THE UNIVERSITY OF ALBERTA

INVESTIGATIONS INTO PROTEIN STRUCTURE, FUNCTION AND *DE NOVO* DESIGN USING A SYNTHETIC PEPTIDE APPROACH

By

Brian P. Tripet ©

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH IN

PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF BIOCHEMISTRY

EDMONTON, ALBERTA

FALL 1999

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-46936-0

Canadä

THE UNIVERSITY OF ALBERTA

LIBRARY RELEASE FORM

NAME OF AUTHOR: Brian P. Tripet

TITLE OF THESIS: Investigations into protein structure, function and *de novo* design using a synthetic peptide approach.

DEGREE: Doctor of Philosophy

YEAR THIS DEGREE GRANTED: 1999

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as herein before provided neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's written permission.

Brian Tripet

304-2 Michener Park
Edmonton, Alberta
Canada
T6H 5B5

Date: Aug 4 1999

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read , and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled "INVESTIGATIONS INTO PROTEIN STRUCTURE, FUNCTION AND *DE NOVO* DESIGN USING A SYNTHETIC PEPTIDE APPROACH" submitted by Brian P. Tripet in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Robert S. Hodges
(Supervisor)

Robert O. Ryan
Robert Ryan

Mark Glover
Mark Glover

David Wishart
David Wishart

Robert Fletterick
(External Examiner)

Date: July 15, 1999

Abstract

Synthetic peptides are powerful tools to enhance our understanding of the structure, function, folding, stability and interactions of proteins. In the present thesis, we have employed the use of several synthetic peptides to extend our present understanding in four major areas of protein research: (1) muscle regulation, (2) motor motility, (3) coiled-coils (4) and protein *de novo* design.

Skeletal muscle regulation is carried out in part by the regulatory protein complex of troponin (Tn) which consists of the three protein subunits: troponin I (TnI), troponin C (TnC) and troponin T (TnT). The Ca^{2+}-dependent change in interactions between the TnI and TnC proteins is regarded as the initial step in the regulatory process. Using several synthetic peptides representing the N and C termini of TnI, we have been able to show that residues 115-131 of TnI bind specifically to the regulatory N-domain of TnC in a Ca^{2+}-dependent manner. Additionally, we showed that residues located C-terminal to the inhibitory region of TnI (residues 140-148) bind to the actin-Tm thin filament and are critical for inducing an inhibitory response similar to that of native TnI. Based on these results, we have proposed a more detailed model of the organization and change in interactions of the troponin complex during the Ca^{2+}-dependent regulatory process.

Kinesin is an intracellular motor protein which utilizes the free energy of ATP hydrolysis to transport membrane bound organelles and vesicles along microtubules. Because of its smaller size in comparison to that of myosin, it has become an ideal model system for investigating the structural changes involved in the mechanism of motor protein motility. Using several synthetic peptides which encompass the kinesin neck region, we demonstrated the presence, location and stability of a parallel two-stranded coiled-coil dimerization domain. Additionally, we showed that the presence of the two helix capping residues Leu 335 and Thr 336 located at the N-terminus of the coiled-coil can further modulate the stability of the coiled-coil domain and hence allude to a mechanism whereby

conformational changes within the motor domains can regulate the opening and closing of the coiled-coil domain during processive stepping along the microtubule.

To understand more clearly the contribution that different amino acid residues can make towards the folding and stability of a protein structure, we systematically substituted 20 different amino acid residues in the hydrophobic core "**a**" and "**d**" positions of a model two-stranded α-helical coiled-coil and determined their resultant effects upon stability and oligomerization state. In general, hydrophobicity was observed to be the dominant factor of stability. However, packing effects for both the hydrophobic as well as polar/charged residues were important for defining the relative stability order and observed effects on oligomerization between the two positions.

As a final test of our understanding of the principles of protein structure and function, we describe the *de novo* design of a heterodimeric coiled-coil dimerization domain which can be used for the rapid detection, purification and characterization of recombinantly expressed peptides and proteins.

Taken together, these studies further our understanding of the principles governing protein structure and function and our ability to design proteins *de novo* .

Acknowledgments

I would like to acknowledge the following individuals who have helped me with this research. First and foremost, I would like to thank my supervisor, Professor Robert S. Hodges who gave me the opportunity to carry out this research in his laboratory. I thank him for his friendship, advice and supervision.

I would like to thank members of the Hodges lab, past and present, who also helped and guided these studies. I thank Jenny Van Eyk for introducing me to muscle regulation. Morris Aarbo for the preparation and purification of various muscle proteins. Paul Semchuck for help with peptide synthesis. Lorne Burke for many odds and ends. Colin Mant and Kurt Wagschal for many helpful discussions, and members of the Alberta Peptide Institute for [illegible] acid and sequencing analysis. I would like to thank Pierre Lavigne for many words of inspiration and help with understanding thermodynamics and molecular modeling.

I would like pay special thank to two members of Dr. Kay's lab. In particular Kim Oikawa for the huge number of CD scans carried out for me, and Les Hicks for the many analytical ultracentrifugation sedimentation equilibrium runs.

This research has been supported by the Alberta Heritage Foundation for Medical Research in the form of studentships and research allowances from the Medical Research Council of Canada, Group in Protein Structure and Function.

PUBLICATIONS
(incurred during the period of this thesis)

B. Tripet, N. Zhou, F. Sonnichsen and R. S. Hodges. Hybritein 1: A *de novo* designed protein displaying cooperative interactions between a metal-ion binding loop and a helix-helix dimerization motif. In Peptides, 1994 (Proceedings of the 23 rd European Peptide Symposium).

B.Tripet, L. Yu, D. Bautista, W.Y.Wong, R.T. Irvin and R. S. Hodges. Engineering a *de novo*-designed coiled-coil heterodimerization domain for the rapid detection, purification and characterization of recombinantly expressed peptides and proteins. *Protein Eng.* 1996, **10**, 101-114.

B.Tripet, J. Van Eyk and R.S Hodges. Mapping of a second actin-tropomyosin and a second troponin C binding site within the C-terminus of troponin I and their importance in the Ca^{2+} -dependent regulation of muscle contraction. *J. Mol . Biol.* 1996, **271**, 728-750.

B.Tripet, R.D. Vale and R.S Hodges. Demonstration of coiled-coil interaction within the kinesin neck region using synthetic peptides: Implication for motor activity. *J. Biol. Chem.* 1996, **272**, 8946-8956.

J.E. Van Eyk, L.T. Thomas, **B.Tripet,** R.J. Wiesner, J.R. Pearlstone, C.S. Farah, F.C. Reinach and R.S. Hodges. Distinct TnI regions regulate Ca^{2+}-dependent potentiation and Ca^{2+}-sensitivity of the acto-S1-TM ATPase activity of the thin filament. *J.Biol. Chem.* 1996, **272**, 10529-10537.

A. P. Campbell, D. L. Bautista, **B. Tripet,** W.Y. Wong, R.T. Irvin, R.S. Hodges and B.D. Sykes. Solution Secondary structure of a bacterially expressed peptide from the receptor binding domain of Pseudomonas aeruginosa pili strain PAK: A heteronuclear multidimensional NMR study. *Biochemistry* 1997, **36**, 12791-12801.

R.T. McKay, **B. Tripet**, R.S. Hodges and B.D. Sykes. Interaction of the second binding region of troponin I with the regulatory domain of skeletal muscle troponin C as determined by NMR spectroscopy. *J.Biol.Chem.* 1997, **272**, 28494-28500.

Chong, P., Chan, N., Kandil, A., **Tripet, B.**, James, O., Yang, Y. P., Shi, S. P. and Klein, M. A strategy for rational design of fully synthetic glycopeptide conjugate vaccines. *Infect Immun* 1997, **65**, 4918-25.

S.T. Kwok, **B.Tripet**, J.H. Mann, M.S. Chana, P. Lavigne, C.T. Mant and R.S. Hodges. Structural cassette mutagenesis in a de novo designed protein: proof of a novel concept for examining protein folding and stability. *Biopolymers, Peptide Sci.* 1998, **47**, 101-123.

K. Wagschal, **B.Tripet**, and R.S. Hodges. *De novo* design of a model peptide sequence to examine the effects of a single amino acid substitutions in the hydrophobic core on both stability and oligomerization state of coiled-coils. *J. Mol. Biol.* 1999, **285**, 785-803.

R.T. McKay, **B. Tripet**, J.R. Pearlstone, L.B. Smillie and B.D. Sykes. Defining the region of Troponin-I that binds to TnC. *Biochemistry*. 1999, **38**, 5478-5489.

P.M. Okamoto, **B.Tripet**, J.Litowski, R.S. Hodges and R.B. Vallee. Multiple coiled-coils are involved in Dynamin self-assembly. *J. Biol..Chem.* 1999, **274**, 10,277-10286.

D.E. Culham, **B.Tripet,** K.I. Racher, R.T. Voegele, R.S. Hodges and J.M. Wood. The role of the carboxyl terminal α-helical coiled-coil domain in osmosensing by transporter ProP of *E. Coli* , *J. Mol. Biol.* 1999, (submitted).

K.Wagschal, **B.Tripet**, P. Lavigne, C.T. Mant and R.S. Hodges. Effects of substituting 20 different amino acid residues in the hydrophobic core "a" position of a *de novo* designed coiled-coil on the stability and oligomerization state. 1999, (submitted).

M. Chana, **B. Tripet** and R. S. Hodges, Analysis of the Coiled-Coil regions within Several Kinesin-like Neck Regions 1999, (in preparation).

B. Tripet, S. Grothe, G. De Crescenzo, M. O'Connor-McCourt, and R. S. Hodges Kinetic analysis of the interactions between Troponin C and Troponin I binding peptides: A surface plasmon resonance and fluorescence study. 1999, (in preparation).

B.Tripet, K. Wagschal, P. Lavigne and R.S. Hodges. The Effect of Side-Chain Characteristics on the Stability and Oligomerization State of a *De Novo* Designed Model Coiled-Coil: Twenty amino acid substitutions in a "d" position. 1999, (in preparation).

B.Tripet, M.Chana and R.S. Hodges. STABLECOIL: Using stability data derived from model synthetic *de novo* designed coiled-coils to predict the location and stability of native coiled-coils 1999 (in preparation).

CONFERENCES AND ABSTRACTS

1. The Fifteenth International Symposium on Protein, Peptide and Polynucleotide Analysis, Boston Massachusetts Nov. 18-21, 1995. A 30 minute presentation describing the development and application of the use of *de novo* designed coiled-coils in purifying recombinant peptides and proteins.

2. Eleventh Symposium of the Protein Society, Boston Massachusetts July 12-16, 1997. Poster presentation "Engineering a de novo-designed coiled-coil heterodimerization domain for the rapid detection, purification and characterization of recombinantly expressed peptides and proteins.

3. Molecular Biophysics of the Cytoskeleton: Formation, Structure, Function and Interactions Aug. 18-22, 1997, Banff, Alberta, Canada. Presentation "Analysis of the stability of de novo designed coiled-coils enhances our understanding of motor and filament protein interactions.

4. First International Peptide Symposium, Kyoto, Japan Nov. 30- Dec. 5, 1997.

5. First Peptide Engineering Meeting, Osaka National Research Institute, Osaka, Japan Dec. 7-9, 1997.

6. The 3rd Australian Peptide Conference, Laguna Quays, The Whitsundays, Queensland, Australia, Oct. 4-9, 1998. Two posters entitled "Biosensor Studies using a Novel Heterodimerization Domain for Peptide-Ligand Display: Troponin C-Troponin I Peptide Interactions" and "Structural Investigation of the Coiled-Coil Neck Regions of Kinesin and Kinesin-Like Proteins: Implications for Motor Structure and Function" were presented.

TABLE OF CONTENTS

Page

Chapter I.

A. Introduction

1. General Overview 1
2. Muscle regulation 1
3. Molecular motility: the kinesins 13
4. Coiled-coils 18
5. The *de novo* design of an affinity purification tag 22

B. References 22

Chapter II. Materials and Methods

A. Materials 29

B. Methods

Peptide synthesis, cleavage and purification 29
Amino acid analysis 33
Mass spectrometry 33
Muscle proteins 34
Preparation of peptides, proteins and complexes for ATPase assays 34
Acto-S1-Tm ATPase Assays 35
Ca^{2+} sensitivity assays 36
Rp$_{1\text{-}40}$ competition assays 37
Actin-Tm centrifugation studies 37
TnC-TnI peptide affinity chromatography 38
Circular dichroism measurements 38
Calculation of denaturation midpoints, free energy of unfolding ($\Delta G_u^{H_20}$) and differences in the free energies of unfolding ($\Delta\Delta G_u$) 40
Size exclusion chromatography with laser light scattering 41
Preparation of oxidized peptides 41
Sedimentation equilibrium and velocity studies 42
Helical propensity/hydrophobicity analysis 42
High performance size exclusion chromatography of the model coiled-coil peptides 44

RP-HPLC analysis of the model a and d peptide analogs 45
Coiled-coil prediction analysis 46
Recombinant engineering (enzymes, oligomers and general techniques) 46
Bacterial strains and plasmids 46
Synthetic E-coil gene 47
Plasmid construction 47
Transfer and screening 48
Construction of the PAK-pilin gene fusion 48
Construction of the PAK-pilin(met) gene fusion 49
Growth, induction and lysis of bacterial containing recombinant plasmids 49
Preparation of the biotinylated K-coil 50
Preparation of the peptide affinity matrix 50
Preparation of the affinity column 51
Purification of recombinant peptides 51
RP-HPLC analysis of the affinity purification method 52
Purification from whole cell lysate 53
SDS-PAGE 53
Western blotting using anti-PAK-pilin 53
Western blotting using biotinylated K-coil 54
CNBr digest of PAK-pilin(met)-E-coil 55
N-terminal sequence analysis of *E.coli* derived peptides 55
C. References 56

Chapter III. Mapping of a Second Actin-Tropomyosin and a Second Troponin C Binding Site within the C Terminus of Troponin I, and their Importance in the Ca^{2+} -dependent Regulation of Muscle Contraction.

A. Introduction 59
B. Results
Identifying the critical residues within TnI_{96-148} involved in inhibition .. 63
Actin-Tm centrifugation studies 68
Neutralization of the inhibitory activity of TnI and TnI peptides by TnC (plus Ca^{2+}) 71
Effect of T·C complex on the inhibitory activity of TnI and TnI

peptides (plus Ca^{2+}) 75
Neutralization of the inhibitory activity of TnI and TnI peptides (minus Ca^{2+}) 79
Affinity chromatography of the TnI peptides 79
Ca^{2+}-sensitivity 82
Competition between the N and C-terminal regions of TnI 85

C. Discussion

Inhibitory sites of interaction 88
The importance of residues 116-131 for TnC and T·C neutralization 90
Binding of the inhibitory region to the central helix and C-domain 92
Structural/functional role of the N terminus of TnI (Rp_{1-40}) 93
Transmission of the calcium signal to TnT-Tm 94
A model for the Ca^{2+}-dependent regulation of muscle contraction 95
D. Conclusion 98
E. Related research 99
F. References 100

Chapter IV. Demonstration of Coiled-coil Interactions within the Kinesin Neck Region Using Synthetic Peptides: Implication for motor activity.

A. Introduction 109

B. Results

CD analysis of the kinesin neck peptides 112
Oxidation of Cys_{330} 117
Stability and concentration dependence of the kinesin peptides 119
Size-exclusion chromatography with laser light scattering 122
Destabilizing effects of Tyr_{344}, Glu_{347} and Asn_{351} in the hydrophobic core 122
Prediction of helical propensity and hydrophobicity in the kinesin neck . 127

C. Discussion

Distinct subdomains of the kinesin neck 130
Coiled-coil interactions within the dimerization domain are organized in a strong-weak-strong manner 135
A model for kinesin motility involving structural transitions within the neck region 136
D. References 140

Chapter V. Investigation of the structural features within the human kinesin neck region.

A. Introduction 144

B. Results

Coiled-coil prediction and analysis of the human kinesin neck region 147

Contribution of the N- and C-terminal heptads to stability and dimerization of the human kinesin coiled-coil 151

Kinesin neck coiled-coil interactions with the N-terminal β-linker region 158

Effect of helix capping interactions between the β-linker and the coiled-coil regions on the stability of the kinesin neck coiled-coil 161

Examination of the potential role Trp 340 in unfolding the coiled-coil domain ? 164

C. Discussion

Importance of the N and C terminal heptads of the kinesin neck coiled-coil........ 166

β-linker and coiled-coil interactions 169

Does the kinesin neck coiled-coil open or not ? 172

Is helix capping a unique structural feature within conventional kinesins only ? 174

A new model for kinesin processivity 176

D. Conclusions 178

E. References 179

Chapter VI. *De novo* design of a model peptide sequence to examine the effects of single amino acid substitutions in the hydrophobic core on both stability and oligomerization state of coiled-coils

A. Introduction 183

B. Results and discussion.

Peptide design 185

Choice of analogs 192

CD spectroscopy 193

Side-chain hydrophobicity 196

Protein stability 196

Oligomerization state 198
Quaternary structure determination 203
Relationship between protein concentration and stability 205
HPSEC under denaturing conditions 208
Comparison of protein denaturation obtained from CD spectroscopy and HPSEC 213
Oligomerization state control by disulfide bridge placement in the hydrophobic core 214
C. Conclusions 216
D. References 217

Chapter VII. The role of position "a" in determining the stability and oligomerization state of α-helical coiled-coils: 20 amino acid stability coefficients in the hydrophobic core of proteins

A. Introduction 225
B. Results and discussion.
Peptide design 228
CD spectroscopy 231
Side-chain hydrophobicity 239
Protein stability 243
Protein stability versus natural occurrence at position "a" 256
Oligomerization state 258
C. Conclusions 264
D. References 265

Chapter VIII. The effect of side-chain characteristics on the stability and oligomerization state of a *de Novo* designed model coiled-coil: Twenty amino acid substitutions in a "d" position.

A. Introduction 271
B. Results
De novo design of the model "d" coiled-coil sequence 274
Secondary structure analysis 278
GdnHCl denaturation studies of the model "d" analogs 281
Hydrophobicity 283
Oligomerization 285
Trimer to monomer switch 291

Calculation of the relative stability difference ($\Delta\Delta G_u$ of unfolding) between model "d" analogs and their comparison to hydrophobicity...... 293
Stability versus hydrophobicity 294
Structural analysis of Met, Trp and Asp 298
Protein stability versus natural occurrence at position "d" 302
Comparison of position "d" versus "a" stability data. 306
C. Discussion
Secondary structure, hydrophobicity and packing 314
Comparison of the stability results 316
Effects of substitutions on the *m*-values 318
Oligomerization 319
D. Conclusions 321
E. References 322

Chapter IX. Engineering a *de novo*-designed coiled-coil heterodimerization domain for the rapid detection, purification and characterization of recombinantly expressed peptides and proteins

A. Introduction 326
B. Results
General concept 329
De novo design of the coiled-coil sequence 332
Preferential formation of the heterodimer 336
Construction of the E-coil fusion vector pRLDE 336
Preparation of the affinity matrix 339
Production and purification of the E-coil and PAK-pilin-E-coil fusion peptide 341
Purification from whole-cell lysate extract 346
SDS-PAGE analysis of the affinity purified peptides 348
Detection of the affinity purified PAK-pilin-E-coil peptide 351
Specificity of the biotinylated K-coil 352
CNBr digest 352
C. Discussion
Advantages of *de novo* design 353
Structural advantages of the coiled-coil heterodimerization domain 353
Detection method 354

Elution conditions 355
Immobilized ligand column 356
Other potential applications 357
Application to the purification of isotopically labeled peptides 358
Openness to redesign 358
D. Conclusions 359
E. Related research 360
F. References 360

Chapter X. Future experiments 366

LIST OF TABLES

Table **Page**

III-1 Amino acid sequences of the synthetic troponin I peptides 65

III-2 Affinity chromatography of the TnI peptides 81

III-3 Titration of TnC and T·C complex with calcium 84

IV-1 Ellipticities and stabilities of the synthetic kinesin peptides 116

V-1 Ellipticities and stabilities of the kinesin peptides 155

VI-1 CD spectroscopy results for coiled-coil X19a analogs 194

VI-2 Biophysical data for the model X19a coiled-coil analogs 195

VI-3 Determination of peptide asymmetry by analytical ultracentrifugation 204

VII-1 CD spectroscopy data of the 20 model X19a analogs 232

VII-2 Side-chain hydrophobicity data determined from amphipathic
α-helical peptides, random coil peptides and N^{α}-acetyl amino acids 237

VII-3 Stability and oligomerization state data for the 20 X19a analogs 246

VIII-1 Biophysical characterization of the synthetic model "d" peptides 279

VIII-2 Oligomerization data 288

VIII-3 Summary of stability and statistical occurrence data 304

VIII-4 Estimation of packing effects for amino acid side-chains in the
"a" and "d" positions of a two-stranded α-helical coiled-coil............. 309

IX-1 Sequences of oligonucleotides used in the construction of the
expression vector pRLDE, the PAK-pilin and PAK-pilin(met) genes 331

IX-2 Amino acid sequences of synthetic and recombinant peptides 335

LIST OF FIGURES

Figure		Page
I-1	Organization and proposed sliding filament model of muscle contraction ...	3
I-2	Schematic diagram of the myosin molecule	4
I-3	Schematic diagram of the thin filament proteins	6
I-4	Crystal structures of TnC using MOLSCRIPT	7
I-5	Schematic representation of the various structural and functional regions within troponin I and T proteins	10
I-6	Simplistic scheme showing the change in direction and interactions of the N-terminal myosin head with the thin filament during the ATP hydrolysis cycle	12
I-7	Schematic diagram of the domain organization of the kinesin molecule	16
I-8	Coiled-coil characteristics	19
I-9	Ribbon representation of several coiled-coil containing proteins	20
II-1	Schematic diagram depicting the various steps carried out during peptide synthesis	30
II-2	Structures of the *in situ* coupling agents used for peptide synthesis	31
II-3	Activation of amino acids using *in situ* coupling reagents	32
III-1	Inhibition of the acto-S1-Tm ATPase activity by TnI and TnI peptides	66
III-2	Actin-Tm centrifugation studies	69
III-3	Neutralization of TnI or TnI peptide inhibition by TnC (plus Ca^{2+})	72
III-4	Neutralization of TnI or TnI peptide inhibition by T·C complex (plus Ca^{2+})	76
III-5	Neutralization of TnI or TnI peptide inhibition by TnC (minus Ca^{2+})	78
III-6	Ca^{2+} dependent release of TnI peptides 96-115, 96-131 96-148 and TnI ..	83

III-7 Competitive displacement of TnI or TnI peptides from TnC (plus Ca^{2+}) by Rp_{1-40} 86

III-8 Schematic summary of the interactions of TnI residues 1 to 40 and 96 to 148 with TnC and actin-Tm, and their role in the Ca^{2+}-dependent regulation of muscle contraction 97

IV-1 Amino acid sequence of the kinesin peptides used in this study 113

IV-2 CD spectra of the kinesin peptides 114

IV-3 CD spectra of the K2 peptide in the reduced and oxidized state, as well as in the presence of 50% TFE 118

IV-4 GdnHCl denaturation and concentration dependence profiles of kinesin peptides 120

IV-5 SEC of the kinesin peptides 123

IV-6 Effects of substituting Tyr^{344}, Glu^{347} and Asn^{351} 126

IV-7 Helical propensity and hydrophobicity of the kinesin neck region 129

IV-8 End and side views of a two-stranded α-helical coiled-coil model from residues 330 to 369 of the kinesin neck region 134

IV-9 Schematic representation of the possible conformational changes occurring in the neck region during kinesin movement 139

V-1. Prediction of a-helical coiled-coil formation within the human kinesin neck region (residues 300-400) 149

V-2 Amino acid sequences of the synthetic peptides used in this study 152

V-3 CD and GdnHCl profiles of K1b and K3b peptides 154

V-4 CD and GdnHCl profiles of K2, K2b and K2c peptides 159

V-5 Molecular interactions occurring between Leu 335, Thr 336, Gln 339, and Trp 340 at the N-terminus of the kinesin neck region.......... 163

V-6 Substitution of Leu 335, Thr 336 and Trp 340 residues 165

V-7 Organization and possible scenarios for the helix capping residues in processivity 175

VI-1 The amino acid sequence and nomenclature of the Leu, Ile, Asn, Ala and Glu model X19a analogs examined in this study 186

VI-2 Proposed models of the interacting peptides in an oxidized and reduced state 187

VI-3 End view of proposed X19a three-stranded reduced coiled-coil from the N-terminus 189

VI-4 End and side views of the structure proposed for X19a in and oxidized, two-stranded monomeric oligomerization state 190

VI-5 GdnHCl denaturation profiles for the analogs monitored by change in molar ellipticity using CD or retention time by SEC 197

VI-6 Sedimentation equilibrium analysis of N19a(ox) and L19a(ox) 200

VI-7 Sedimentation equilibrium analysis of N19a(cm) and L19a(cm) 201

VI-8 HPSEC of the Ile19a analog after equilibration at 500 μM, 170 μM and 2.5 μM 206

VI-9 GdnHCl denaturation profiles of reduced and oxidized peptide Ile19a at peptide concentrations 50 μM and 2.5 μM 207

VI-10 HPSEC of the Asn19a(ox) peptide at 2.5 μM under denaturing conditions ranging from 0.4 M GdnHCl to 3 M GdnHCl 210

VI-11 HPSEC of the Ile19a(ox) peptide at 2.5 μM under denaturing conditions ranging from 0.4 M GdnHCl to 4 M GdnHCl 212

VI-12 Sedimentation equilibrium analysis of C36dL19a(ox) and C36dL19a(red) 215

VII-1 Amino acid sequence of the 76-residue disulfide bridged model coiled-coil X19a 229

VII-2 Proposed models of the interchain disulfide-bridged peptides 230

VII-3 Reversed-phase high performance liquid chromatographic separation of each analog to determine relative side-chain hydrophobicity at pH 7 236

VII-4 Correlation of side-chain hydrophobicity data 241

VII-5 Chemical denaturation profiles 244

VII-6 Comparison of chemical denaturation midpoints and relative stabilities obtained using GdnHCl and Urea 249

VII-7 Correlation of side-chain hydrophobicity and relative stability 252

VII-8 Comparison of side-chain occurrence with relative side-chain stability 257

VII-9 Molecular modeling of the Leu and Trp analogs as two and three-stranded coiled-coils 263

VIII-1 Model "d" coiled-coil sequence 276

VIII-2 GdnHCl denaturation profiles of the model "d" peptides 282

VIII-3 Correlation plot of the relative RP-HPLC retention times of the monomeric α-helix obtained from the model "d" peptide analogs versus those obtained from the model "a" peptide sequence 284

VIII-4 Representative HPLC-size exclusion chromatograms for the model "d" peptide analogs Glu, Ser and Val 287

VIII-5 Sedimentation equilibrium analysis of model "d" analogs Val, Ser and Lys. 292

VIII-6 Comparison of the relative stability of each substituted amino acid with its relative hydrophobicity for the "d" position analogs 297

VIII-7 Molecular models depicting the sidechain orientation of Trp, Met, Gln and Asp residues in the "d" position of a two-stranded parallel coiled-coil 300

VIII-8 Comparison of the relative stability contribution each amino acid imparts to the hydrophobic core "a" or "d" positions 303

VIII-9 Comparison of the frequency of occurrence of each amino acid residue in an "a" or "d" position of the heptad repeat in native coiled-coil sequences 308

VIII-10 Comparison of the statistical occurrence of each amino acid residue occurring in an "a" or "d" position of the heptad repeat in native coiled-coil sequences 313

IX-1 Schematic representation of the cloning, detection and purification of an E-coil-tagged peptide or protein using the coiled-coil fusion system 330

IX-2 Helical wheel representation of one heptad in three possible coiled-coils 334

IX-3 pRLDE expression vector and restriction sites 338

IX-4 Schematic representation of the steps involved in the formation of the K-coil affinity matrix 340

IX-5 RP-HPLC analysis of the eluates from each step during affinity purification of recombinantly expressed E-coil and PAK-pilin-E-coil peptides 343

IX-6 RP-HPLC analysis of recovered E-coil and PAK-pilin-E-coil peptides after affinity purification of whole-cell extracts 347

IX-7 SDS-PAGE analysis of the affinity purification of E-coil and PAK-pilin-E-coil peptides from *E.coli* cells 349

IX-8 Immunodetection of recombinant PAK-pilin-E-coil using PK99H or biotinylated K-coil 350

X-1 Schematic representation of the steps involved in utilizing the E- and K-coiled-coil template as a peptide ligand display system for surface plasmon resonance analysis.. 367

LIST OF ABBREVIATIONS

Amino acids:

Ala, A	Alanine
Arg, R	Arginine
Asn, N	Asparagine
Asp, D	Aspartic acid
Cys, C	Cysteine
Gln, Q	Glutamine
Glu, E	Glutamic acid
Gly, G	Glycine
His, H	Histidine
Ile, I	Isoleucine
Leu, L	Leucine
Lys, K	Lysine
Met, M	Methionine
Nle,	Norleucine
Phe, F	Phenylalanine
Pro, P	Proline
Ser, S	Serine
Thr, T	Threonine
Trp, W	Tryptophan
Tyr, Y	Tyrosine
Val, V	Valine

Ac-	Acetylated N-terminus
acto-S1	Actin and myosin subfragment 1 complex
A1/A2	Myosin associated light chains
ADP	Adenosine 5'-diphosphate
ATP	Adenosine 5'-triphosphate
-Amide	Amidated C-terminus
BB	Benzylbenzoic acid
Boc or t-Boc	tert-Butyloxycarbonyl
BSA	Bovine serum albumin
CD	Circular dichroism
CNBr	Cyanogen bromide
-COOH/-OH	Free carboxyl terminus
DCM	Dichloromethane
DIEA	N,N-Diisopropylethylamine
DMF	N,N-Dimethylformamide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EDT	1,2,-Ethandithiol
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethyleneglycobis(β-aminoethylether)N,N,N',N',-tetraacetic acid
F	Fluorescence
F-actin	Fibrous actin protein
GdnHCl	Guanidine hydrochloride
$[\text{GdnHCl}]_{1/2}$	Guanidine hydrochloride denaturation midpoint
HBTU	2-(1H-Benzotriazol-1-yl)-1,1,3,3-tetrmethyluronium hexafluorophosphate
HEPES	N-2-hydroxyethylpiperazine-N'2-ethansulfonic acid

HF	Hydrofluoric acid
HPLC	High performance liquid chromatography
HPSEC	High performance size-exclusion chromatography
I.D.	Internal diameter
Ip	TnI inhibitory peptide residues 104-115
IPTG	Isoproppyl-b-D-thiogalactopyranoside
K_d	Apparent equilibrium dissociation constant
m	Slope of the free energy versus denaturant concentration plots
MW	Molecular weight
NMR	Nuclear magnetic resonance
NH_2-	Free amino terminus
PAK	Pilin strain PAK of pseudomonas aeruginosa bacteria
PBS	Phosphate buffer saline
Pi	Inorganic phosphate
PCR	Polymerase chain reaction
pCa	-log[Ca2+]
Rp	TnI N-terminal regulatory peptide residues 1-40
RP-HPLC	Reversed-phase high-performance liguid chromatography
SEC	Size-exclusion chromatography
S1	Myosin subfragment 1
SDS-PAGE	Sodium dodececylsulfate polyacrylamide gel electrophoresis
SE	Sedimentation Equilibrium analysis
Tn	Troponin
TnI	Troponin I
TnC	Troponin C
TnT	Troponin T
Tm	Tropomyosin

Tris	tris-(hydroxymethyl)amino methane
TFA	Trifluoroacetic acid
TFE	Trifluoroethanol

Chapter I

A. Introduction

I.1 General Overview

The present thesis is submitted in 'paper format' as outlined in the graduate student thesis handbook. Each chapter includes an introduction, a results section, a general discussion, and a conclusion. For enhanced readability all methods pertaining to each individual chapter have been grouped together and are presented in Chapter II as "Material and Methods". Because several of the chapters deal with differing topic areas, a general introduction reviewing each particular area is presented below, along with the sought goal or aim. In general, this thesis although wide in scope deals with studies involved in understanding better the molecular determinants of protein structure, function, stability and interactions.

It is important to note that due to the magnitude of the third part of this work (i.e., the study of the 20 amino acid substitutions in a model coiled-coil protein **a** and **d** position) the project was split into two parts, with first authorship given to K. Wagschal, while the second part (Chapter VIII), once fully revised, will be published by B.Tripet as first author.

I.2 Muscle regulation (Part 1)

Movement has always been associated with life and the elucidation of the molecular mechanisms underlying such processes have been the subject of several research groups for several decades. It is hoped that a sound understanding of this field will not only advance our understanding of contractile and motile processes, but also our understanding of protein-protein interactions, conformational changes, Ca^{2+} signaling and chemical to mechanical energy conversions in general.

As the first study presented in this thesis deals specifically with aspects of muscle regulation, a brief review of the various proteins involved in skeletal muscle contraction and regulation are described below. For more detailed information concerning aspects of the regulatory process, please consult either the more extensive introduction in Chapter III or the following excellent recent reviews: Farah and Reinach, 1995; Gergely, 1998; Reinach *et al.*, 1997; Squire and Morris, 1998; Tobacman 1996; Zot and Potter, 1987.

The thick and thin filaments.

The basic structural organization of vertebrate striated muscle tissue consists of two types of filaments, the thick and thin, which by electron microscopy appear as two interdigitated lines which slide past one another when muscle cells are triggered to contract (Huxley, 1957, 1969). A schematic representation of this organization is shown in Figure I-1. The thick filament is composed largely of dimeric myosin molecules which are organized in a bipolar manner in which their C-terminal α-helical coiled-coil tails are oligomerized to form the core of the thick filament, while their globular N-terminal heads project out at both ends (Figure I-1 and I-2). Functional characterization of myosin has shown that the ATPase activity and actin-binding properties are located specifically within the globular N-terminal domain head. Hence it is presently believed that it is the cyclical interaction of these heads with the thin filament and their concomitant hydrolysis of ATP that provides the driving force for the sliding of the thick filament over the thin filament and thus muscle contraction.

Related specifically to the studies described herein, the myosin molecule can be cleaved into sub-fragments through exposure to certain enzymes (Figure I-2). For example, chymotrypsin cleaves the heavy chain of myosin in the neck region between the coiled-coil stalk and the globular head domain to yield a single-headed sub-fragment termed S1 which in context with the associated essential light chains A1 (MW ~21,000) and A2 (MW

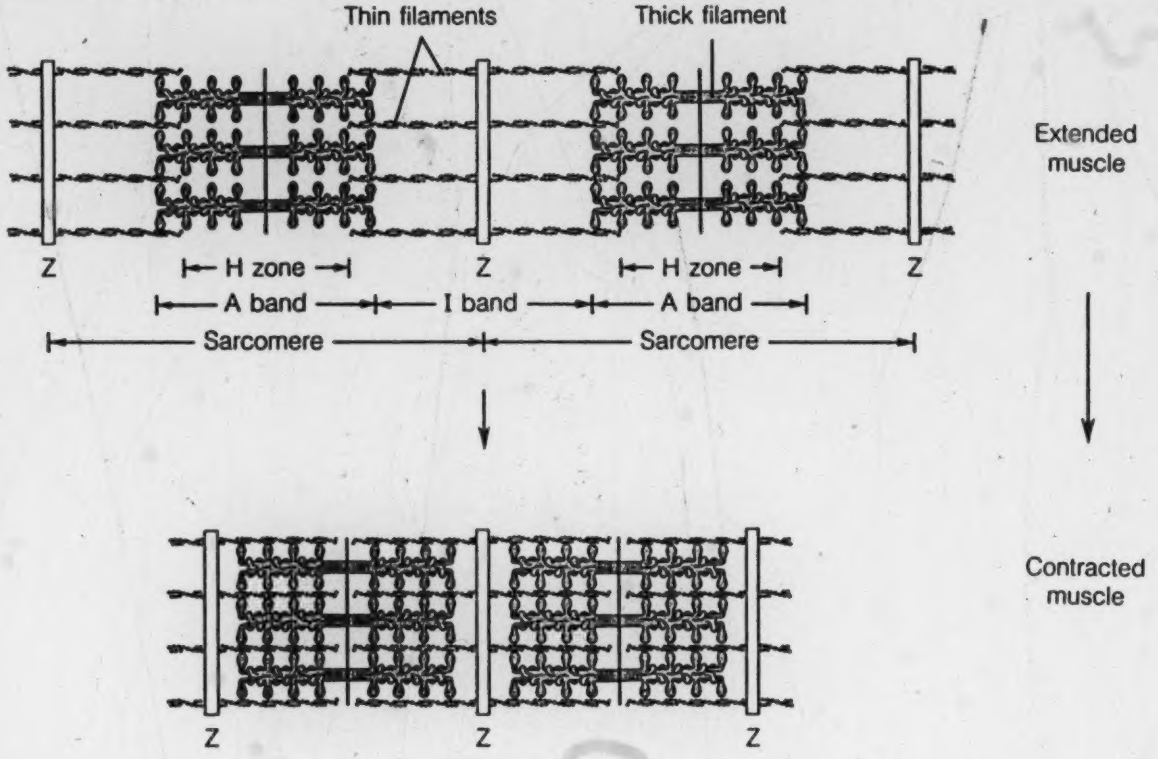

Thin filaments
Thick filament
Extended muscle
Z
H zone
Z
H zone
Z
A band
I band
A band
Sarcomere
Sarcomere
Contracted muscle
Z
Z
Z

Figure I-1. Organization and proposed sliding filament model of muscle contraction. Shown are two sarcomeres which are extended and contracted. The location of the thick and thin filaments are denoted in the figure. Adapted from Mathews and vanHolde, 1990.

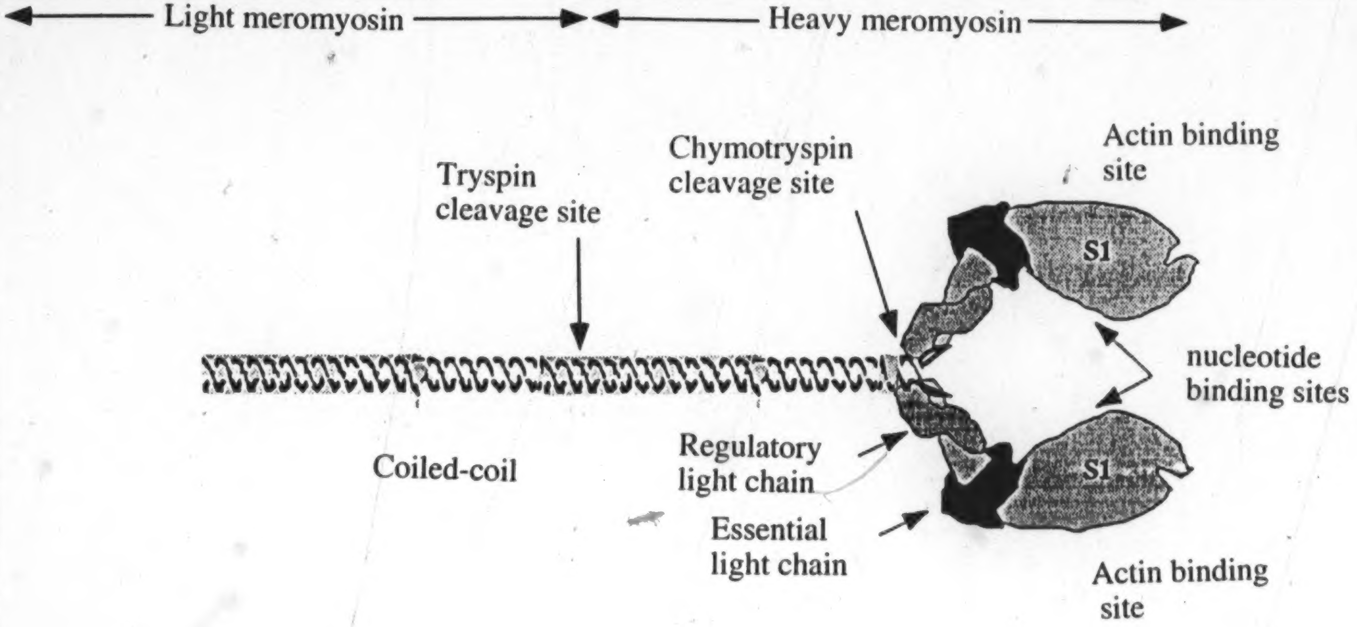

Light meromyosin
Heavy meromyosin
Tryspin
cleavage site
Chymotryspin
cleavage site
Actin binding
site
S1
nucleotide
binding sites
S1
Coiled-coil
Regulatory
light chain
Essential
light chain
Actin binding
site

Figure I-2. Schematic diagram of myosin. Shown are two heavy chains of myosin which form globular heads at their N-terminus (containing the actin and nucleotide binding sites) and wind around one another at their C-terminus to form an α-helical coiled-coil. The essential and regulatory light chains are indicated. Myosin may be enzymatically cleaved with trypsin or chymotrypsin to form cleavage products which contain largely the globular heads and associated light chains e.g. chymotrypsin cleavage produces the globular heads with associated light chains termed S1 and no coiled-coil region. This was the molecule used in the studies described in this thesis.

~16,200) retains the complete ATPase properties of the molecule and can be used as a mimic of the thick filament in *in vitro* solution assays (Weeds and Taylor, 1975).

The thin filament on the other hand is made up of the three major proteins: actin, tropomyosin (Tm) and troponin (Tn). A schematic representation of this structure is shown in Figure I-3. The backbone of the thin filament consists of two strands of polymerized actin monomers coiled about one another in a helical array. Lying in each grove is one molecule of tropomyosin, a long two-stranded α-helical coiled-coil protein, which lies end to end and overlap slightly. Each tropomyosin molecule spans ~380 Å along the actin filament , and thus creates a periodic repeat every seven actin monomers. Associated at approximately the center of each tropomyosin molecule is one molecule of troponin (Tn).

The troponin complex.

In 1963 Ebashi demonstrated that the Ca^{2+}-dependent control of muscle contraction was mediated via the two thin filament associated proteins tropomyosin and troponin. Subsequent studies showed that troponin actually consists of three non-identical protein subunits which have been named troponin C (TnC), troponin I (TnI) and troponin T (TnT) based on their apparent functional properties. For the purpose of this introduction the structural and functional properties of each individual subunit of troponin is described further below.

TnC is the calcium binding subunit, which for rabbit skeletal muscle, is composed of 159 residues with a calculated molecular weight of 18,081 Da . TnC has a pI value in the range of 4.1-4.4 due to a high content of glutamic and aspartic acid residues. Of the three troponin subunits, TnC is the only protein whose atomic structure has been solved to date by both X-ray crystallography and NMR spectroscopy (Gagne, 1995; Houdusse *et al.*, 1997; Herzberg and James, 1985; Slupsky and Sykes, 1995; Slupsky *et al.*, 1995; Strynadka *et al.*, 1997). Elucidation of the initial crystal structures of turkey and avian skeletal TnC showed that TnC exists as a dumbbell-shaped molecule with two globular

The Thin Filament

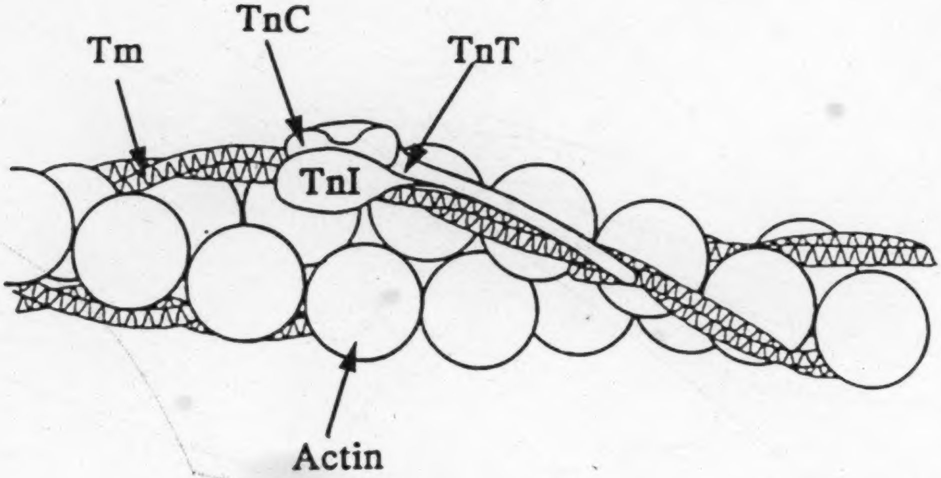


Figure I-3. Schematic diagram of the thin filament proteins. The actin filament is composed of two strands of polymerized actin monomers. Tropomyosin (Tm), polymerized head to tail lies near the grooves of the actin filament. One tropomyosin molecule spans seven actin monomers. The troponin complex interacts with both actin and tropomyosin through troponin I (TnI) and troponin T (TnT) proteins, respectively. Troponin C (TnC) is the calcium binding subunit of the troponin complex. The troponin complex together with tropomyosin forms the basic regulatory unit of the thin filament. The figure is adapted from Heeley *et. al.* 1987.

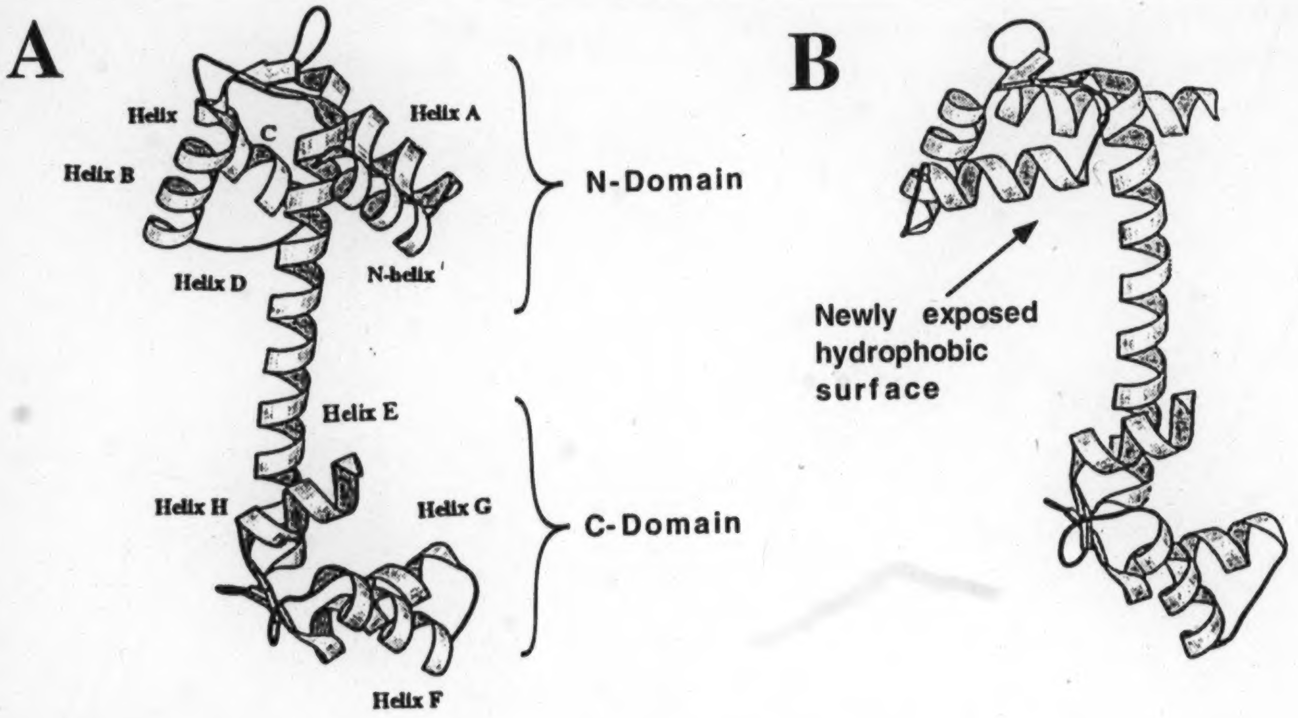

A
Helix
C
Helix A
Helix B
Helix D
N-helix
Helix E
Helix H
Helix G
Helix F
N-Domain
C-Domain
B
Newly exposed
hydrophobic
surface

Figure I-4. Crystal structures of TnC presented using MOLSCRIPT. **(A)** depicts the overall fold of TnC when only the C-terminal binding sites are filled with calcium (2-Ca^{2+}-TnC). **(B)** depicts the overall fold of TnC when all four of the calcium binding sites are filled with calcium (4-Ca^{2+}-TnC). The nine helices are denoted N, A-H, respectively. The hydrophobic surface area exposed upon binding Ca^{2+} in the N-terminal domain is indicated with an arrow.

domains termed the N- and C-domains, respectively (Fig. I-4). Each domain contains two metal ion-binding sites of the EF-hand structural motif (Kretsinger, 1973) which are numbered I-IV according to their observed order in the primary sequence. In skeletal TnC, the two Ca^{2+} binding sites in the C-domain (sites III and IV) bind Ca^{2+} with high affinity ($K_a \sim 10^7$ M^{-1}) as well as Mg^{2+} ions competitively ($K_a \sim 10^3$ M^{-1}) (Potter and Gergely, 1974). The two binding sites within the N-domain (sites I and II) bind Ca^{2+} specifically but with lower affinity (Ka $\sim 10^5$ M^{-1}) (Potter and Gergely, 1974). The binding of Ca^{2+} to the N-terminal domain sites (I and II) have been implicated as the relevant triggering sites for initiating muscle contraction. Hence these sites have also been termed the "regulatory sites' (Johnson *et al.*, 1979; Robertson *et al.*, 1979; Farah and Reinach 1995 and references there in).

On the basis of extensive similarities in amino acid sequence between the N- and C-domains and on the assumption that the structure of the Ca^{2+} saturated N-domain would be similar to that of the C-domain, Herzberg, Moult and James (1986) proposed a model for the Ca^{2+} saturated state of TnC. The authors proposed that upon binding Ca^{2+} the N, B and C helices (denoted A-H according to the initial crystallographic analysis) would open up (relative to helices A and D) to expose a new "hydrophobic surface area " or "pocket" which would be critical for relieving the inhibition of the actomyosin ATPase by TnI (Fig. I-4). Recently this model has received strong experimental support due in part to the X-ray crystallography and NMR structures of the Ca^{2+} saturated N-domain of TnC (Gagne, 1995; Houdusse *et al.*, 1997; Slupsky and Sykes, 1995; Slupsky *et al.*, 1995; Strynadka *et al.*, 1997). However, the exact residues of TnI which bind to this site are presently unknown.

Troponin I (TnI) is the inhibitory subunit of the troponin complex. TnI can by itself bind to actin filaments and inhibit the actomyosin ATPase activity (Hartshorne and Muller, 1968; Schaub and Perry, 1969), however, this inhibitory activity can be greatly enhanced by the addition of tropomyosin even at ionic strengths where tropomyosin neither binds nor inhibits actomyosin (Eisenberg and Kielley 1974; Eaton, 1975; Talbot and Hodges, 1979). The region of TnI responsible for its inhibitory activity was first localized by Syska *et al.*, (1976) to a CNBr cleavage fragment (CN4) corresponding to residues 96-116. Subsequent studies using synthetic peptides have since shown that residues 104-115 represents the minimal sequence necessary for inhibition (Talbot and Hodges, 1981). Based on its ability to bind the actin-Tm filament and inhibit the acto-myosin ATPase activity, and be released by TnC in a Ca^{2+}-dependent manner, it is believed that this region of TnI is the triggering switch between muscle contraction and relaxation states (Van Eyk *et al.*, 1993).

Other notable regions of TnI include residues 1-47 (Figure. I-5) which bind TnC in the presence of Ca^{2+} and Mg^{2+} ions (Syska *et al.*, 1976). Structural and functional studies have shown that this region interacts primarily with the C-domain of TnC and is important for maintaining the structural integrity of the troponin complex in the absence of Ca^{2+} ions (Farah *et al.*, 1994; Vassylyev *et al.*, 1998). The region corresponding to residues 40-96 is also interesting in that it contains a periodic hydrophobic repeat similar to that observed in the two-stranded α-helical coiled-coil molecule of tropomyosin (Pearlstone and Smillie, 1985). Thus it has been proposed that this region may form a coiled-coil interaction with a similar repeating pattern observed in the TnT molecule (residues 197-250) (Pearlstone and Smillie, 1985), although no studies have been reported to date to support such a hypothesis. Presently little is known about the functional importance of residues C-terminal to the inhibitory region (residues 115-181) and thus are explored herein.

Troponin T (TnT), the last of the troponin subunits, contains 259 amino acid residues and has been found to play the primary role of linking the troponin complex to tropomyosin. TnT is a highly polar molecule with acidic sidechains near its amino terminus (1-39) and basic sidechain residues near its C-terminus (221-259). Antibody experiments by Ohtsuki (1974, 1975 and 1979) have clearly demonstrated that the molecule is organized into two separate domains which extend lengthwise along tropomyosin (Figure I-3). The N-terminal region, termed T1 (residues 1-159), interacts with tropomyosin at the head-to-

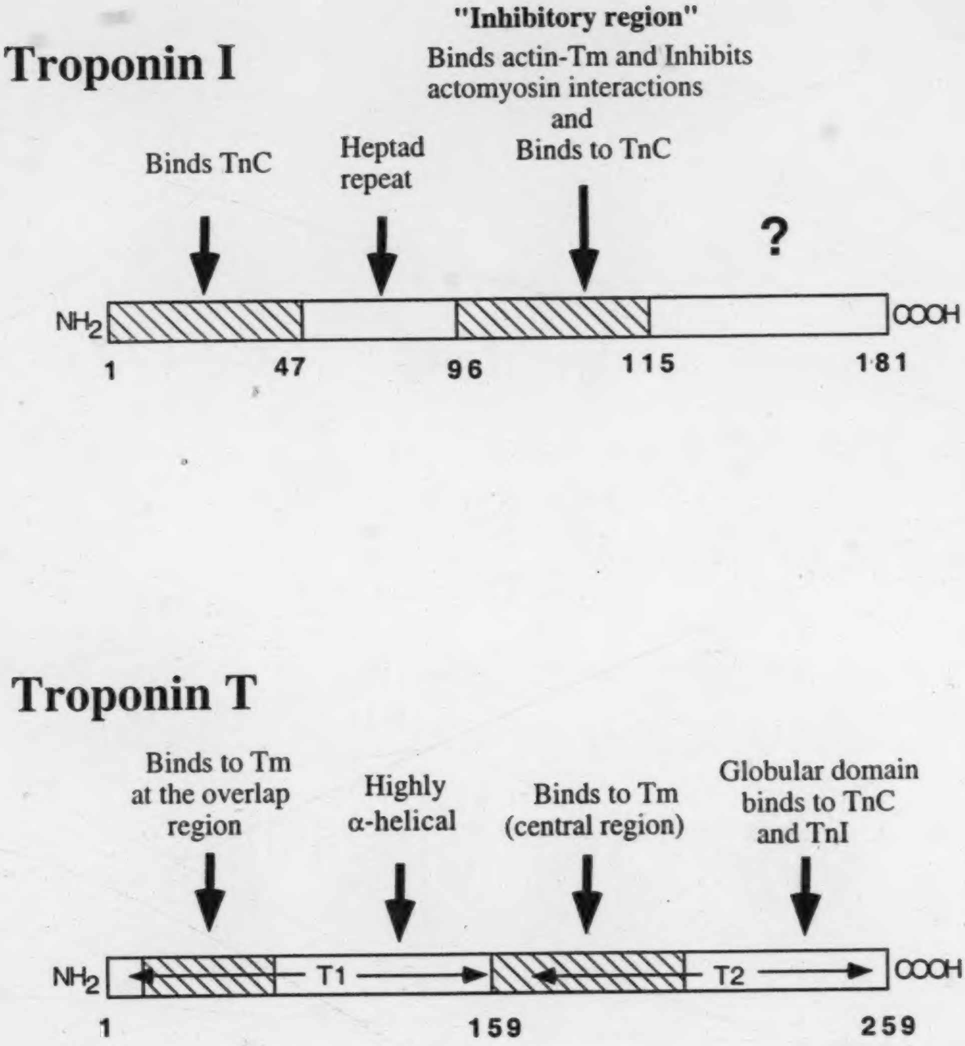

Troponin I
"Inhibitory region"
Binds actin-Tm and Inhibits actomyosin interactions and Binds to TnC
Binds TnC
Heptad repeat
?
NH2
COOH
1
47
96
115
181
Troponin T
Binds to Tm at the overlap region
Highly α-helical
Binds to Tm (central region)
Globular domain binds to TnC and TnI
NH2
T1
T2
COOH
1
159
259

Figure I-5. Schematic representation of the various structural and function regions identified within troponin I and T proteins.

tail overlap region while the C-terminal region, termed T2 (residues 160-259), interact with tropomyosin (as well as TnI and TnC) at its centre (Figure I-5). Although TnI can inhibit the actomyosin ATPase activity and TnC releases this inhibition, full Ca^{2+}-dependent control occurs only in the presence of TnT. Thus it is believed that TnT plays an important structural/functional role in relaying the structural changes within the troponin complex to Tm and the actin filament. At present no high resolution X-ray crystallography or NMR structure exist for any part of TnT alone nor in a binary or tertiary interaction with any of the other troponin proteins.

Regulation of the thick and thin filaments.

The troponin complex in conjunction with tropomyosin and Ca^{2+} regulate the interactions of the thick and thin filaments. The most widely accepted model for this processes involves a combination of the steric and allosteric blocking models first proposed by Eisenberg *et al.*, 1980; Geeves *et al.*, 1984; Haselgrove, 1972; Huxley, 1972; Parry and Squire, 1973. In this model (depicted in Figure I-6) the heavy chain heads of myosin can bind to actin in two states (states iii and iv) which are each assumed to have different mechanical properties. The initial cross-bridge attachment between actin and myosin is weak (iii), however, upon phosphate release, there is a transition to a strong bound state . The transition from the weak binding state to the strong binding state of actin-myosin corresponds to an ~ 90^{o} to 45^{o} movement of the myosin head (termed the power stroke). In the presence of low Ca^{2+} concentrations (10^{-7} M) Tm-Tn is capable of blocking the movement of the myosin head from 90^{o} to 45^{o} by adopting an "off" position on the thin filament facilitated by the binding of TnI to actin. In the presence of high Ca^{2+} concentrations (10^{-5} M), Ca^{2+} binding to the TnC subunit releases the interactions of the TnI protein with actin and subsequently causes Tm to occupy a deeper position in the actin groove. It is presently believed that the movement of Tm also causes the concomitant

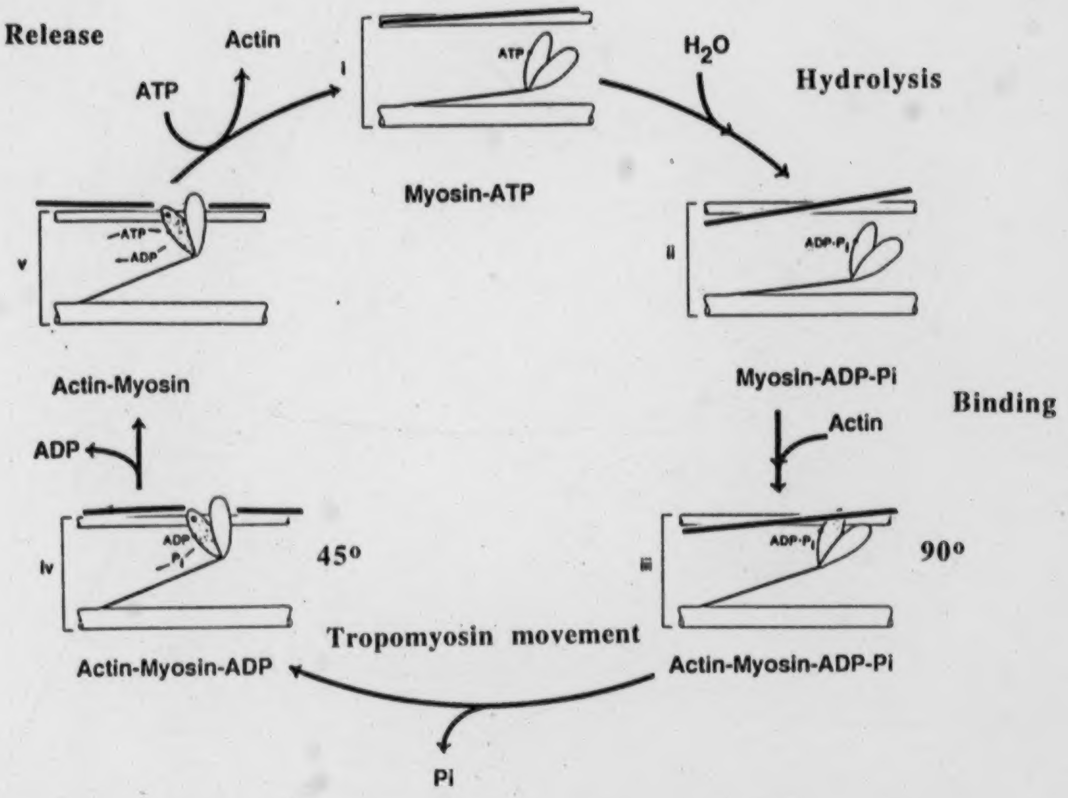

Release
ATP
Actin
i
ATP
H2O
Hydrolysis
Myosin-ATP
v
—ATP—
—ADP
ii
ADP·Pi
Actin-Myosin
Myosin-ADP-Pi
Binding
Actin
ADP
iv
ADP
—Pi
45°
iii
ADP·Pi
90°
Tropomyosin movement
Actin-Myosin-ADP
Actin-Myosin-ADP-Pi
Pi

Figure I-6. Simplistic scheme of the coupling of ATP hydrolysis cycle and Tm movement to steps involved during muscle contraction (see text for more details). Figure adapted from Eisenberg *et al.*, (1980).

conformational change in the actin monomers to allow myosin to progress to its 90° state (iv). This model appears to be in agreement with S1 binding in the absence of Ca^{2+} and the observation of potentiation of the acto-S1-ATPase rate (ATPase rate greater than actin and S1 alone) indicative of different conformational states of actin.

Aim of the first study

In spite of the extensive number of studies carried out to date, several issues still remain in regards to clearly understanding at a molecular level the Ca^{2+}-dependent change in interactions between TnI and TnC and their ability to regulate the position of Tm on the actin filament.

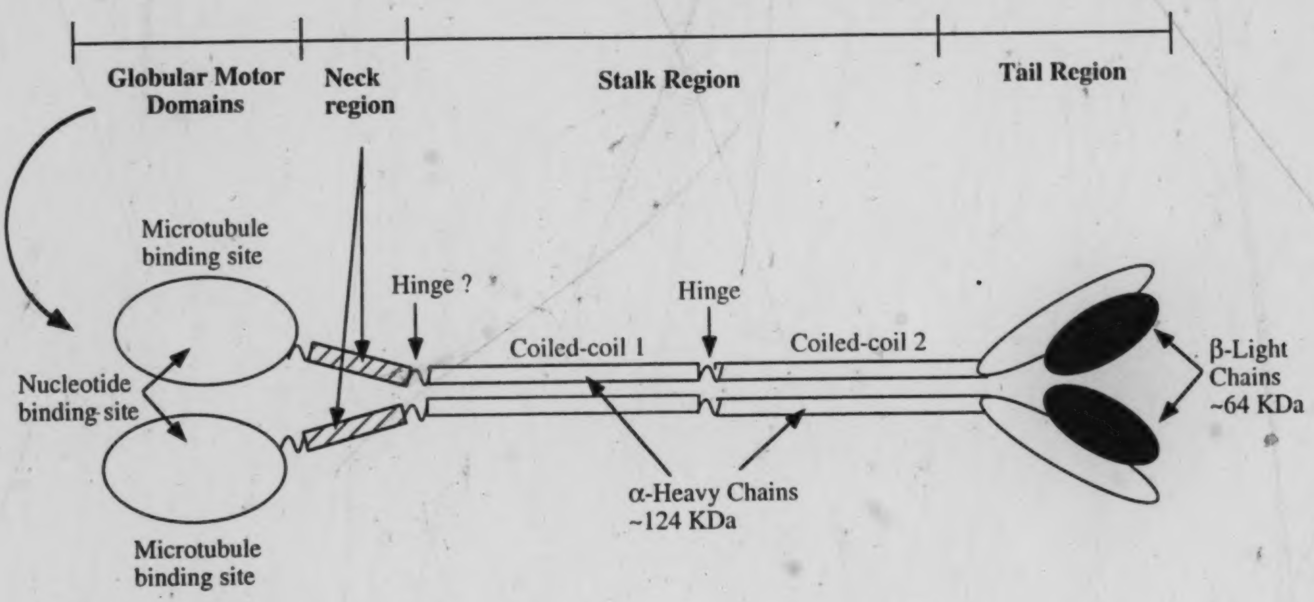
Firstly, although the TnI inhibitory region (residues 96-115) can bind actin-Tm and inhibit the acto-myosin ATPase (Syska *et al.*, 1976), and be neutralized by TnC in the presence of Ca^{2+}, the preferred binding site on TnC (derived from a number of photochemical X-linking and proteolytic mapping studies) is suggested to be within the C-domain of TnC. Thus other residues within the TnI protein must interact with the N-terminal regulatory domain. Secondly, although residues 96-115 of TnI can induce a similar inhibitory activity to that of TnI, it does so only at a significantly greater molar ratio (~2.5-3.0) than that of intact TnI (Talbot and Hodges, 1979; Van Eyk 1993, 1997). Thus this result would suggest that other residues within the TnI molecule are also important for mediating the inhibitory activity. Interestingly, recent studies by Farah *et al.*, (1994) and Van Eyk *et al.*, (1997) have indicated such residues may be located within the C-terminal region of TnI (residues 115-148). Thus to explore these issues further, we have prepared several synthetic peptides encompassing various regions within the C-terminal region 96-148 of rabbit skeletal TnI and analyzed each of these peptides in reconstituted thin filament inhibitory and release assays (see chapter III).

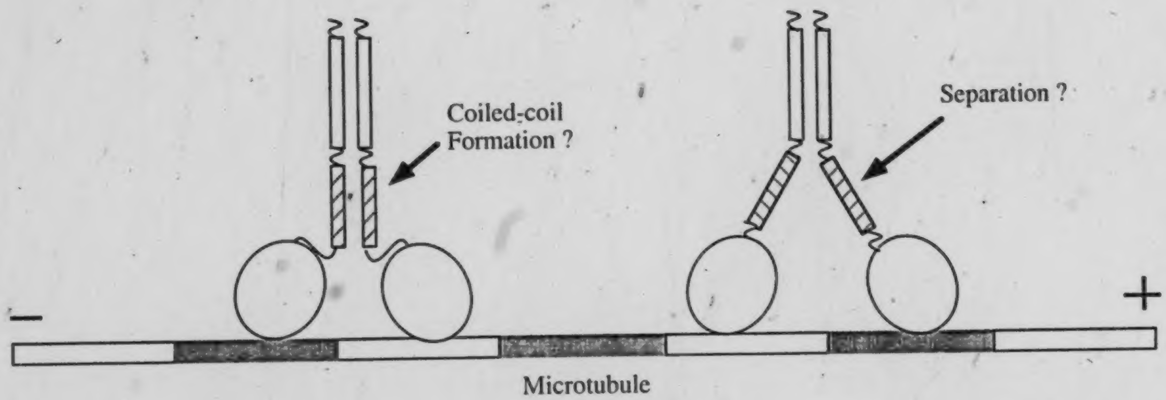
I-3. Molecular motility, the kinesins (Part 2)

Motor proteins are mechanical enzymes that translate the energy of nucleotide hydrolysis into directed movement. The best known motors are those of the myosin superfamily which move along actin filaments (Pollard *et al.*, 1991; Titus, 1993; Baker and Titus, 1998), and briefly we described the properties of the heavy chain myosin molecule of vertebrate skeletal muscle and its interactions with the thin filament above. The second best known family of motor proteins are those of the kinesins. The kinesins encompass a large and ever expanding group of related motor proteins which all contain the presence of a highly conserved ~315 amino acid residue catalytic domain within their polypeptide chain. In contrast to the myosin motors which move along actin filaments, the catalytic domain of kinesins interact specifically with microtubule filaments. Moreover, as opposed to being involved in macroscopic movement, for example as with skeletal muscle contraction, the kinesins are involved in intracellular movements such as cell division, organelle movement, vesicle transport, and general cellular organization (Moore and Endow, 1996; Vale and Fletterick, 1997). For more information on the phylogenics, roles and structures of the molecules within the kinesin superfamily please consult the following excellent reviews: Hirokawa, 1997; Hirokawa, 1998; Hirokawa, *et al.*, 1998; Howard, 1997; Mandelkow and Johnson, 1998; Moore and Endow, 1996; Vale and Fletterick, 1997.

Of the large number of proteins existing within the superfamily, kinesin (also referred to as conventional kinesin) is the best studied member of the superfamily and as such is the only member mentioned hereafter. Conventional kinesin purified from variety of tissue sources exists as an $\alpha_2\beta_2$ heterotetramer, in which two α subunits (heavy chains) and two β subunits (light chains) associate to form a highly elongated molecule with globular termini (Bloom *et al.*, 1988; Kuznetsov *et al.*, 1988; Figure I-7). The kinesin heavy chains are organized into four domains (listed from N to C-terminus): 1) a ~325 amino acid residue globular motor domain head that contains the ATP and microtubule binding sites,

Figure I-7. Top, schematic diagram of the domain organization of the kinesin molecule. Open circles represent the globular N-terminal motor domain heads (residues 1-320) which contain the nucleotide and microtubule binding sites. Residues ~320-370 shown as hatched rectangles denote the 'neck region'. Extending towards the C-terminus (from the neck region) is the "stalk region " (open rectangles) which forms a long two stranded α-helical coiled-coil structure broken at its centre by a proline containing region (termed the hinge). Located at the very C-terminus are the two kinesin light chains (β-chains, black spheres) which associate with the two heavy chains through a coiled-coil interaction and together form the cargo binding site. **Bottom,** schematic diagram showing the kinesin molecule traveling along a microtubule filament. Depicted in the figure are the main questions investigated in the present studies: 1) does the kinesin neck region form a two-stranded α-helical coiled-coil ? and 2) if so, does the neck region contain unique structural features within it that allow it to temporarily unfold during stepping to the next adjacent microtubule binding site.


Globular Motor Domains
Neck region
Stalk Region
Tail Region
Microtubule binding site
Hinge ?
Hinge
Coiled-coil 1
Coiled-coil 2
Nucleotide binding site
β-Light Chains ~64 KDa
α-Heavy Chains ~124 KDa
Microtubule binding site


Coiled-coil Formation ?
Separation ?
−
+
Microtubule

2) a ~50 amino acid residue region adjacent to the globular motor domain (termed the neck region) that is sufficient for allowing dimerization of the motor domains (Huang *et al.*, 1994) and contains a sequence that is predicted to form an α-helical coiled-coil (Stewart *et al.*, 1993; Huang *et al.*, 1994), 3) a long ~450 amino acid residue α-helical coiled-coil domain, termed the stalk (Stewart *et al.*, 1993), and 4) a small globular C-terminus, termed the tail (Hirokawa *et al.*, 1989; Scholey *et al.*, 1989; Yang *et al.*, 1989; de Cuevas *et al.*, 1992). Flexible "hinge" regions are found between the neck and the stalk and in the center of the stalk. The light chains (β subunits) of kinesin, which are not necessary for force-generation, are associated with the smaller globular C-terminus of the heavy chain and are presently believed to be involved in determining cargo specificity (Hirokawa *et al.*, 1989).

Single molecule motility assays have demonstrated that conventional kinesin is endowed with a unique property in that a single kinesin molecule can move continuously along a microtubule for several microns in a series of 8 nm steps, which corresponds to the distance between tubulin binding sites along the microtubule protofilament (Svoboda *et al.*, 1993). Such processive movement, which is not displayed by muscle myosin or ciliary dynein, very likely represents a specialized adaptation that enables few kinesin motors to transport membrane organelles efficiently within cells. Functional studies on recombinantly expressed kinesin heavy chains have begun to uncover regions that are necessary for kinesin motility. For example bacterial expression of the first 340 amino acids of the *Drosophila* kinesin heavy chain (which contains the core N-terminal globular motor domain and the first ~10 amino acids of the neck) produces a monomeric protein that generates directed motility when many motors are interacting simultaneously with a single microtubule in gliding motility assays (Yang *et al.*, 1990; Stewart *et al.*, 1993). However, these monomeric kinesins do not exhibit processive movement when assayed as individual motors in a single molecule fluorescence motility assay (Vale *et al.*, 1996) or a bead assay (Berliner *et al.*, 1995). A kinesin motor containing the complete motor and neck domains, on the other hand, forms a dimer and also exhibits processive movement (Hackney, 1995;

Vale *et al.*, 1996). Collectively, these studies suggest that the dimeric structure of kinesin is not essential for force-generation *per se*, although it does appear to be required for processive movement. This raises the possibility that processive movement may involve a hand-over-hand coordination of the two kinesin heads.

The proposed α-helical coiled-coil domain in the kinesin neck may be structurally important for coordinating the activities of the two kinesin heads during processive movement. The existence of a coiled-coil structure in close proximity to the motor domains, however, raises important questions concerning its exact boundaries and stability, since the connection between the heads must be sufficiently extensible to allow the two motor domains to span the distance between two tubulin dimers during movement. It is also important to determine the thermodynamic properties of the neck coiled-coil to ascertain if it could partially or totally "un-coil" during the generation of a power stroke.

Thus to address these questions and others we have prepared a series of synthetic peptides corresponding to different regions of the human kinesin neck (residues 316-383) and analyzed each peptide for its respective secondary structure content and stability. These results are presented in chapter IV and V.

I-4. Coiled-coils (Part 3)

The α-helical coiled-coil is a common assembly motif found in a wide variety of proteins in nature. These include cytoskeletal and extracellular matrix components, transcription factors, cellular and viral surface proteins, dyneins and kinesins, muscle proteins, tumor suppressers and oncogene products (Adamson *et al.*, 1993, Cohen and Parry, 1990). Typically, coiled-coils consist of two to five right-handed amphipathic α-helices which coil around each other to form a slight left handed supercoil (Figure I-8 and Figure I-9). At present, the two and three-stranded associative states are the most often observed structures in nature. Due to its structural simplicity, and small size, the coiled-coil has proven to be an excellent model system for not only exploring the determinants for

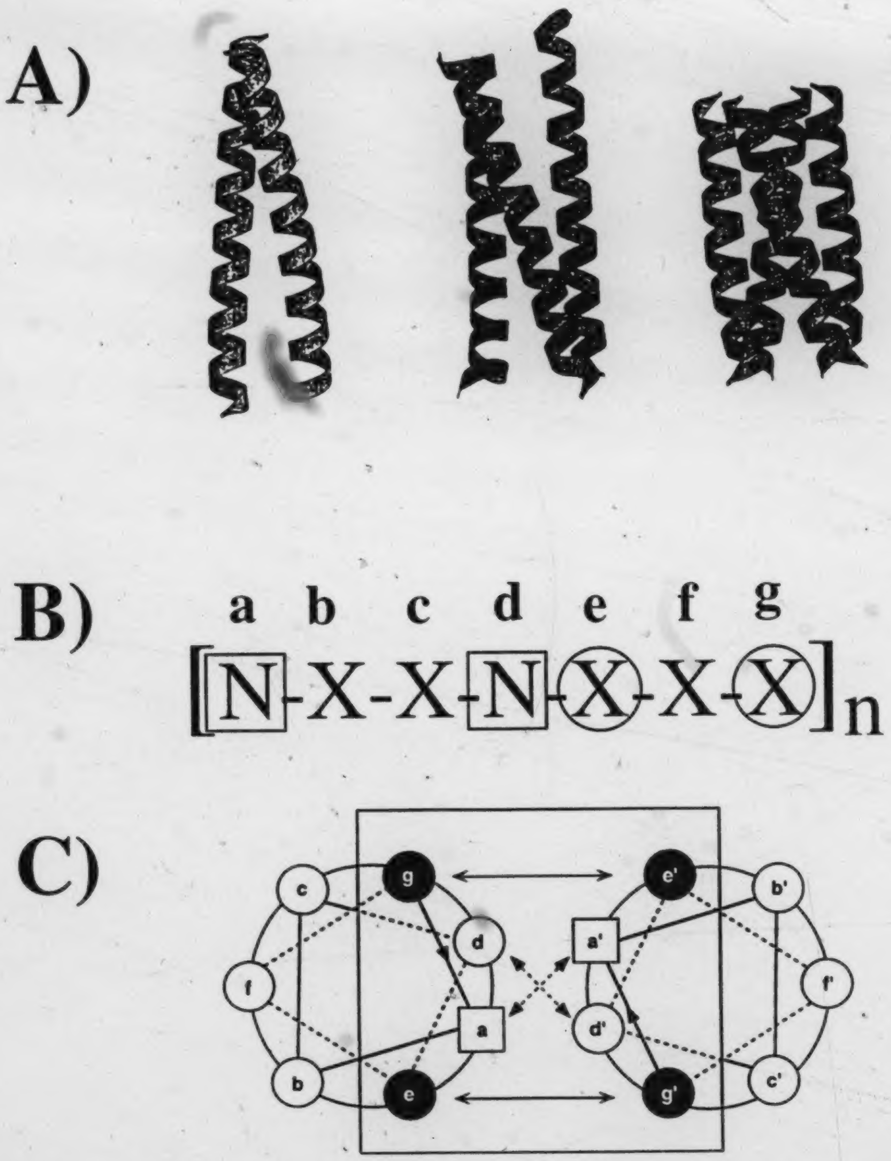
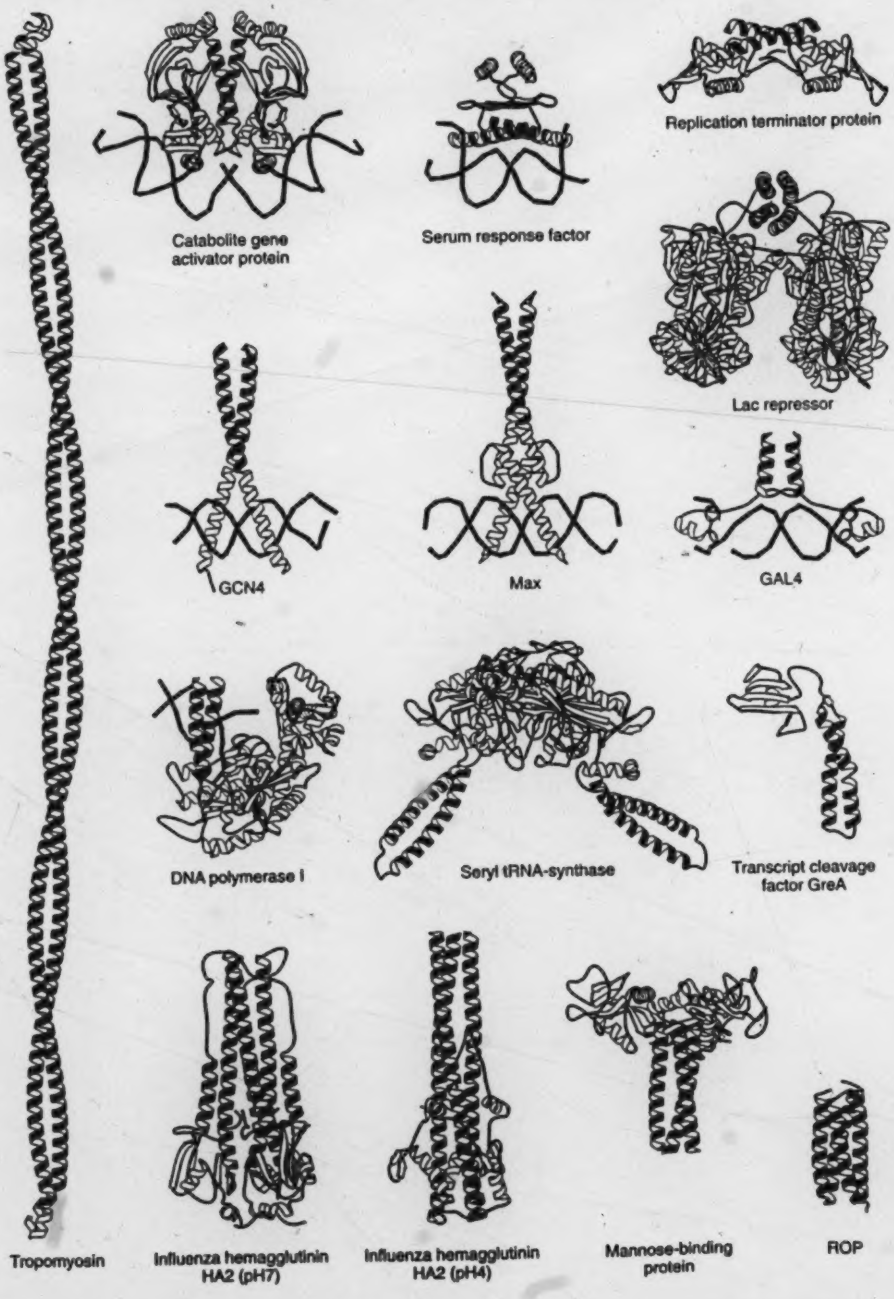


Figure I-8. Coiled-coil characteristics. **A)** Ribbon representation of the most frequently observed coiled-coil oligomerization states (i.e. two, three and four stranded). **B)** The sequences of coiled-coils have a heptad pattern of residues of the form (abcdefg)n where a and d are usually apolar, and positions e and g are frequently charged. **C)** End-on view of two helices showing apolar interactions between residues a and d, and electrostatic interactions between residues e and g. The interfacial residues exist within the box.



Native Coiled-Coil Proteins

Catabolite gene activator protein
Serum response factor
Replication terminator protein
Lac repressor
GCN4
Max
GAL4
DNA polymerase I
Seryl tRNA-synthase
Transcript cleavage factor GreA
Tropomyosin
Influenza hemagglutinin HA2 (pH7)
Influenza hemagglutinin HA2 (pH4)
Mannose-binding protein
ROP

Figure I-9. Ribbon representation of several coiled-coil containing proteins. Figure is adapted from A. Lupas (1996).

coiled-coil formation but also the fundamental features of protein folding and protein-protein interactions. For example, model synthetic coiled-coils have recently been employed to investigate α-helical propensity values (O'Neil and Degrado, 1990), inter- and intrahelical electrostatic interactions (O'Shea, *et al.*, 1993; Kohn *et al.*, 1995a,b; Spek *et al.*, 1998, Yu *et al.*, 1996) helix dipole effects (Kohn *et al.*, 1997), hydrophobicity and packing (Zhou *et al.*, 1992a,b,c; Zhu *et al.*, 1993; Moitra *et al.*, 1997), association kinetics (Chao *et al.*, 1996) heterodimer specificity (Lavigne *et al.*, 1998), oligomerization (Harbury *et al.*, 1993), as well as general principles of protein *de novo* design (Bryson *et al.*, 1995, 1998, Hodges 1996, and Kohn and Hodges 1998).

The sequence of coiled-coils are characterized by a seven residue (heptad) repeat traditionally denoted **(abcdefg)**$_n$ (McLachlan and Stewart 1975) where positions **a** and **d** are primarily occupied by hydrophobic residues (Hodges *et al.*, 1972), positions **e** and **g** by opposite charged residues, and positions **b**, **c** and **f** by polar or charged residues. When this pattern is configured into an α-helix, the hydrophobic side-chains at positions "**a**" and "**d**" create an amphipathic α-helix where the non-polar surfaces of two or more such subunits can oligomerize in a parallel and in register fashion to form a continuos hydrophobic core as a result of the "knobs-into-holes" packing as proposed by Crick (1953). To date, this structure has been verified by greater than 50 different high resolution X-ray and NMR structures deposited in the protein data bank.

Extensive structure function studies investigating the importance of each of the various positions within the heptad repeat have firmly established that the coiled-coil hydrophobic core **a** and **d** positions are in large part responsible for defining the folding and stability of the structure (Betz *et al.*, 1995; Cedervall *et al.*, 1997; Greenfield & Hitchcock-DeGregori, 1995; Harbury *et al.*, 1993; Hodges , 1992; Monera *et al.*, 1996; Zhou *et al.*, 1992a,b; Zhu *et al.*, 1993). Moreover, it has been shown that type and placement of nonpolar, polar or charged residues, e.g., Ile, Asn or Lys, in these positions can results in distinct oligomerization states (Gonzales *et al.*, 1996a,b; Harbury *et al.*, 1993; O'Shea *et al.*, 1991). It follows then that both the *de novo* design of a coiled-coil sequence with a specific stability and oligomerization state, and the identification and preliminary characterization of a coiled-coil structural motif in naturally occurring protein sequences requires knowledge of the contribution that each amino acid residue can make toward these properties when present at an **a** or **d** position. At present, however, no systematic and comprehensive study has yet been carried out using well-defined model coiled-coils sequences to determine the exact contribution that each side-chain in the hydrophobic core can make towards stability and oligomerization state. Thus, in chapters VI, VII and VIII, we describe the *de novo* design and complete biophysical characterization of two model coiled-coil proteins , X19**a** (also termed model **a**) and model **d,** to which we have systematically substituted 19 naturally occurring amino acids (cysteine was omitted) and one non-natural amino acid (Orn) into a central hydrophobic core **a** and **d** position, respectively.

I-5. The *de novo* design of an affinity purification tag (Part 4)

The *de novo* design of model proteins is an important endeavor that not only tests our understanding of protein folding and structure, but also lays the ground work for the design of novel proteins with unique biological activities. The purpose of the final part of this thesis (Chapter IX) was thus to investigate whether we could apply the large number of principles learned herein regarding stability, oligomerization state and specificity of protein-protein interactions to *de novo* design a coiled-coil heterodimerization domain that could act as a mimic to naturally occurring high affinity protein ligand-receptor interactions. This entailed designing *de novo* two separate coiled-coil sequences in which the interfacial residues positions (i.e., position **a**, **d**, **e** and **g**) would prevent homodimerization but would allow folding and high stability in a heterodimeric state. A full description of the *de novo* design process, its utility and general applicability are explored in chapter IX.

B. References

Adamson, J. G., Zhou, N. E. and Hodges, R. S. (1993) *Curr Opin Biotechnol* **4**, 428-437.

Baker J.P., Titus, M.A., (1998) *Curr Opin Cell Biol* , **10**, 80-86

Berliner, E., Young, E. C., Anderson, K., Mahtani, H. and Gelles, J. (1995) *Nature*, **373**: 718- 721.

Betz, S., Fairman, R., O'Neil, K., Lear, J. & DeGrado, W. (1995) *Phil. Trans. R. Soc. London B* **348,** 81-88.

Bloom, G. S., Wagner, M. C., Pfister, K. K., and Brady, S. T. (1988) *Biochemistry*, **27**, 3409-3416.

Bryson, J.W., Betz, S.F., Lu, H.S., Suich, D.J., Zhou, H.X., O'Neil, K.T. and DeGrado, W.F. (1995) *Science*, **270**, 935-941.

Bryson, J.W., Desjarlais, J. R., Handel, T.M. and DeGrado, W.F. (1998) Prot. Sci. **7**, 1404-1414.

Cedervall, T., Johansson, M. U. & Åkerström, B. (1997) *Biochemistry* **36,** 4987-4994.

Chao, H., Houston, M. E., Jr., Grothe, S., Kay, C. M., O'Connor-McCourt, M., Irvin, R. T. and Hodges, R. S. (1996). *Biochemistry* **35**, 12175-85.

Cohen, C. and Parry, D. A. D. (1990) *Proteins: Struct. Funct. Genet.* **7**, 1-15.

Crick, F. H. C. (1953) *Acta Crystallogr.* **6**, 689-698.

de Cuevas, M., Tao, T., and Goldstein, L. S. B. (1992) *J. Cell Biol.* **116**, 957-965.

Draper, M.H., and Hodge, A.J. (1949) *Aust. J. Exptl. Med. Sci.* **27**, 465-503.

Eaton, B.L., Kominz, D.R., and Eisenberg, E. (1975). *Biochemistry*, **14**, 2718-2724.

Ebashi, S., (1963) Nature **200**, 1010.

Ebashi, S., and Kodama, A. (1965) *J. Biochem.* **58**, 107-108.

Eisenberg, E. and Kielly, W.W. (1974) *J. Biol. Chem.* **249**, 4742-4748.

Eisenberg, E., Hill, T.L., and Chen, Y. (1980) *Biophys. J.*, **29**, 195-227.

Farah, C. S. & Reinach, F. C. (1995) *FASEB J.* **9,** 755-767.

Farah, C. S., Miyamoto, C. A., Ramos, C. H., da Silva, A. C., Quaggio, R. B., Fujimori, K., Smillie, L. B., and Reinach, F. C. (1994) *J. Biol. Chem.* **269,** 5230-5240.

Gagné, S.M., Li,M.X., McKay, R.T., and Sykes, B.D. (1998) *Biochem. Cell Biol.* **76,** 302-312.

Gagné, S.M., Tsuda, S., Li, M.X.,Smillie, L.B., and Sykes, B.D. (1995) *Nature Struct. Biol.* **2** 784-789.

Geeves, M.A., Goody, R.S., and Gutfreund, H. (1984) *J. Muscle Res. and Cell Motility* 5, 351-361.

Gergely, J. (1998) *Adv. Exp. Med. Biol.* **453**, 169-176.

Gonzales, L. Jr., Woolfson, D. N. & Alber, T. (1996a) *Nature Struct. Biol.* **3,** 1011-1018.

Gonzales, L. Jr., Brown, R. A., Richardson, D. & Alber, T. (1996b) *Nature Struct. Biol.* **3,** 1002-1009.

Greenfield, N. J. & Hitchcock-DeGregori, S. E. (1995) *Biochemistry* **34,** 16797-16805.

Hackney, D. D. (1995) *Nature* **377**, 448-450.

Harbury, P. B., Zhang, T., Kim, P. S. & Alber, T. (1993) *Science* **262,** 1401- 1407.

Hartshourne, D.J., and Muller, H. (1968) *Biochem. Biophys. Res. Commun.* **31**, 647-653.

Haselgrove, J.C. (1972). *Cold Spring Harbor Symp. Quant. Biol.* **37**, 341-348.

Heeley, D.H., Goloninska, K., and Smillie, L.B. (1987) J. Biol. Chem. 262, 9971-9978

Herzberg, O. & James, M. N. G. (1985). *Nature* **313,** 653-659.

Herzberg, O., Moult, J., and James, M. N. (1986). *J. Biol. Chem.* **261,** 2638-2644.

Hirokawa, N. (1997) *Curr Opin Neurobiol* **7**, 605-614.

Hirokawa, N. (1998) *Science* **279**, 519-526.

Hirokawa, N., Noda, Y. and Okada, Y. (1998) *Curr Opin Cell Biol* **10**, 60-73.

Hirokawa, N., Pfister, K. K., Yorifuji, H., Wagner, M. C., Brady, S. T., and Bloom, G. S. (1989) *Cell*, **56**, 867-878.

Hodges, R. S. (1992) *Current Biology* **2,** 122-124.

Hodges, R. S. (1996) *Biochem Cell Biol* **74**, 133-154.

Hodges, R. S., Sodek, J., Smillie, L. B. and Jurasek, L. (1972) *Cold Spring Harbor Symp. Quant. Biol.* **37**, 299-310.

Houdusse, A., Love, M.L., Dominguez, R., Grabarek, Z., and Cohen, C. (1997) *Curr. Biol.* **5**, 1695-1711

Howard, J. (1997) *Nature* **389**, 561-567.

Huang, T.G., Suhan, J., Hackney, D.D. (1994) *J. Biol. Chem.* **269**, 16502-16507.

Huxley, H.E. (1953) *Biochem. et Biophys. Acta* **12**, 387-394.

Huxley, H.E. (1957) *Prog. Biophys. Chem.* **7**, 255-318.

Huxley, H.E. (1969) *Science* **164**, 1356-1366.

Huxley, H.E. (1972) *Cold Spring Harbor Symp. Quant. Biol.* **37**, 361-369.

Jonhson, J. D, Charlton, S.C., and Potter, J.D. (1979) *J. Biol. Chem.* **254**, 3497-3502.

Kohn, W. D. and Hodges, R. S. (1998) *Trends Biotech.* **16**, 379-389.

Kohn, W. D., Monera, O. D., Kay, C. M. and Hodges, R. S. (1995a). *J Biol Chem* **270**, 25495-506.

Kohn, W. D., Kay, C. M. and Hodges, R. S. (1995b) *Protein Sci* **4**, 237-50.

Kohn, W. D., Kay, C. M. and Hodges, R. S. (1997) *J Pept Sci* **3**, 209-23.

Kretsinger, R. H. & Nockolds, C. E. (1973) *J. Biol. Chem.* **248,** 3313-3326.

Kuznetsov, S. A., Vaisberg, E. A., Shanina, N. A., Magretova, N. N., Chernyak,V.Y., and Gelfand, V. I. (1988) *EMBO J.* **7**, 353-356.

Lavigne, P., Crump, M. P., Gagne, S. M., Hodges, R. S., Kay, C. M. and Sykes, B. D. (1998). *J Mol Biol* **281**, 165-181.

Lupas, A. (1996). *Trends Biochem. Sci.* **21,** 375-382.

Mandelkow, E. and Johnson, K. A. (1998)*Trends Biochem Sci* **23**, 429-433.

McLachlan, A. D. and Stewart, M. (1975). *J Mol Biol* **98**, 293-304.

Moitra, J., Szilak, L., Krylov, D. and Vinson, C. (1997). *Biochemistry* **36**, 12567-73.

Monera, O. D., Sönnichsen, F. D., Hicks, L., Kay, C. M. & Hodges, R. S. (1996) *Protein Eng.* **9,** 353-363.

Moore, J. D. and Endow, S. A. (1996) *Bioessays* **18**, 207-219.

O'Neil, K. and DeGrado, W. F. (1990) *Science* **250**, 646-651.

O'Shea, E. K., Klemm, J. D., Kim, P. S. & Alber, T. (1991) *Science* **254,** 539-544.

O'Shea, E. K., Lumb, K. J. and Kim, P. S. (1993). *Curr. Biol.* **3**, 658-667.

Ohtsuki, I. (1974) *J. Biochem.* **75**, 753-765.

Ohtsuki, I. (1975) *J. Biochem.* **77**, 633-639.

Ohtsuki, I. (1979) *J. Biochem.* **86**, 491-497.

Parry, D.A.D., and Squire, J.M. (1973) *J. Mol. Biol.* **75**, 33-55.

Pearlstone, J.R., and Smillie, L.B. (1985) *Can. J. Biochem.* **63**, 212-218.

Pollard, T.D., Doberstein, S.K. and Zot, H.G., (1991). *Annu. Rev. Physiol,* **53**, 653-681.

Potter, J.D. and Gergely, J. (1974). *Biochemistry* **13**, 2697-2703.

Reinach, F.C., Farah, C.S., Monteiro, P.B., Malnic, B. (1997) *Cell Struct. Funct.* **22,** 219-223.

Robertson, S.P., Johnson, J.D., and Potter, J.D. (1981) *Biophys. J.* **34**, 559-569.

Rosa, G., Szent-Gyorgyi, A., Wycyoff, R.W. (1950) *Exptl. Cell. Res.* **1**,194-205.

Schaub, M.C., and Perry, S.V., (1969) *Biochem. J.* **115**, 993-1004.

Scholey, J. M., Heuser, J., Yang, J. T., and Goldstein, L. S. B. (1989) *Nature*, **338**, 355-357

Slupsky, C.M., and Sykes, B.D (1995) *Biochemistry 34*, 15953-15964.

Slupsky, C.M., Reinach, F.C., Smillie, L.B., and Sykes, B.D. (1995) *Protein Sci.* **4**, 1279-1290.

Spek, E. J., Bui, A. H., Lu, M. and Kallenbach, N. R. (1998) *Protein Sci* **7**, 2431-2437.

Squire, J.M, and Morris, E.P. (1998) *FASEB J.*, **12**, 761-771.

Stewart, R.J., Thaler, J.P., and Glodstein, L.S.B. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 5209-5213

Strynadka, N.C.J., Cherney, M., Sielecki, A.R., Li, M.X., Smillie, L.B., and james, M.N.G. (1997) J. Mol. Biol. **273**, 238-255.

Svoboda, K., Schmidt, C. F., Schnapp, B. J. and Block, S. M. (1993) *Nature* **365**, 721-727.

Syska, H., Wilkinson, J. M., Grand, R. J. A., and Perry, S. V. (1976). *Biochem. J.* **153,** 375-387.

Talbot, J. A., and Hodges, R.S. (1979) *J. Biol. Chem.* **254**, 3720-3723.

Talbot, J. A., and Hodges, R.S. (1981) *J. Biol. Chem.* **256**, 2798-2802.

Titus, M.A. (1993). *Curr. Opin. Cell Biol.* **5**,77-81.

Tobacman, L.S., (1996) *Annu. Rev. Physiol.* **58**, 447-481.

Vale, R. D. and Fletterick, R. J. (1997) *Annu .Rev.Cell Dev.Biol* **13**, 745-777.

Vale, R.D., Funatsu, T., Pierce, D. W., Romberg, L., Harada, Y., and Yanagida, T. (1996) *Nature* **380**: 451-453.

Van Eyk, J. E., Strauss, J. D., Hodges, R. S., and Ruegg, J. C. (1993). *FEBS Lett.* **323,** 223-228.

Van Eyk, J. E., Thomas, L. T., Tripet, B., Wiesner, R. J., Pearlstone, J. R., Farah, C. S., Reinach, F. C., and Hodges, R. S. (1997). *J. Biol. Chem.* **272**, 10529-10537.

Vassylyev, D.G., Takeda, S., Wakatsuki, S., Maeda, K., ans Maeda, Y. (1998) *Proc. Natl. Acad. Sci.* **95**, 4847-4852.

Weeds, A. G. & Taylor, R. S. (1975). *Nature* **257,** 54-56.

Yang, J. T., Laymon, R. A., and Goldstein, L. S. B. (1989) *Cell*, **56**, 879-889

Yang, J. T., Saxton, W. M., Stewart, R. J., Raff, E. C.and Goldstein, L. S. (1990). *Science*, **249**, 42-47.

Yu, Y., Monera, O. D., Hodges, R. S. and Privalov, P. L. (1996). *J .Mol .Biol* . **255**, 367-372.

Zhou, N. E., Kay, C. M. & Hodges, R. S. (1992a). *Biochemistry* **31,** 5739-5746.

Zhou, N. E., Kay, C. M. & Hodges, R. S. (1992b). *J. Biol. Chem.* **267,** 2664-2670.

Zhou, N. E., Zhu, B. Y., Kay, C. M. and Hodges, R. S. (1992c) *Biopolymers* **32**, 419-426.

Zhu, B.Y., Zhou, N. E., Kay, C. M. & Hodges, R. S. (1993). *Protein Sci.* **2,** 383-394.

Zot, A. S. & Potter, J. D. (1987). *Annu. Rev. Biophys. & Biophys. Chem.* **16,** 535-559.

CHAPTER II

MATERIALS AND METHODS

A. Materials

Unless otherwise stated, all chemicals and solvents were reagent grade. Diisopropylethylamine, trifluoroacetic acid, and dichloromethane were distilled prior to use. *t*-Butyloxycarbonyl (*t*-Boc) amino acids and resins were purchased from Novabiochem (San Diego, CA). ATP was purchased from Boehringer Mannheim (Indianapolis, IN). DTT was purchased from ICN Biomedicals, Inc (aurora, OH), and rabbit skeletal muscle (rabbit muscle tissue type I, New Zealand white) was purchased from Pel-Freeze Biologicals (Rogers, AR).

B. Methods

Peptide Synthesis, Cleavage and Purification

Synthetic peptides were prepared by solid-phase synthesis methodology using a 4-benzhydrylamine-hydrochloride (MBHA) resin with conventional *N-t-butyloxycarbonyl* (*t*-Boc) chemistry (Figure II-1, II-2 and II-3). Extensions of the chains was carried out using either: (1) an Applied Biosystems Model 430A peptide synthesizer (Foster City, CA.) as described by Sereda *et al.*, (1994) or (2) polypropylene reaction reservoirs where resin washing and amino acid activation and addition were performed manually. Standard side chain protecting groups were routinely used. Peptides which contained methionine residues, were initially prepared as methionine sulfoxide using *t*-Boc-Met(O). After coupling of the final amino acid, all peptides were N-terminally acetylated using acetic anhydride. Peptidyl-resin was cleaved by reaction with hydrogen fluoride (20 ml/g resin) containing 10% anisole and 1% 1,2-ethanedithiol for 1.5 h at -5 °C. Crude peptides were washed several times with cold ether, extracted from the resin with glacial acetic acid and lyophilized. The crude peptides were purified by RP-HPLC on a SynChropak semi-

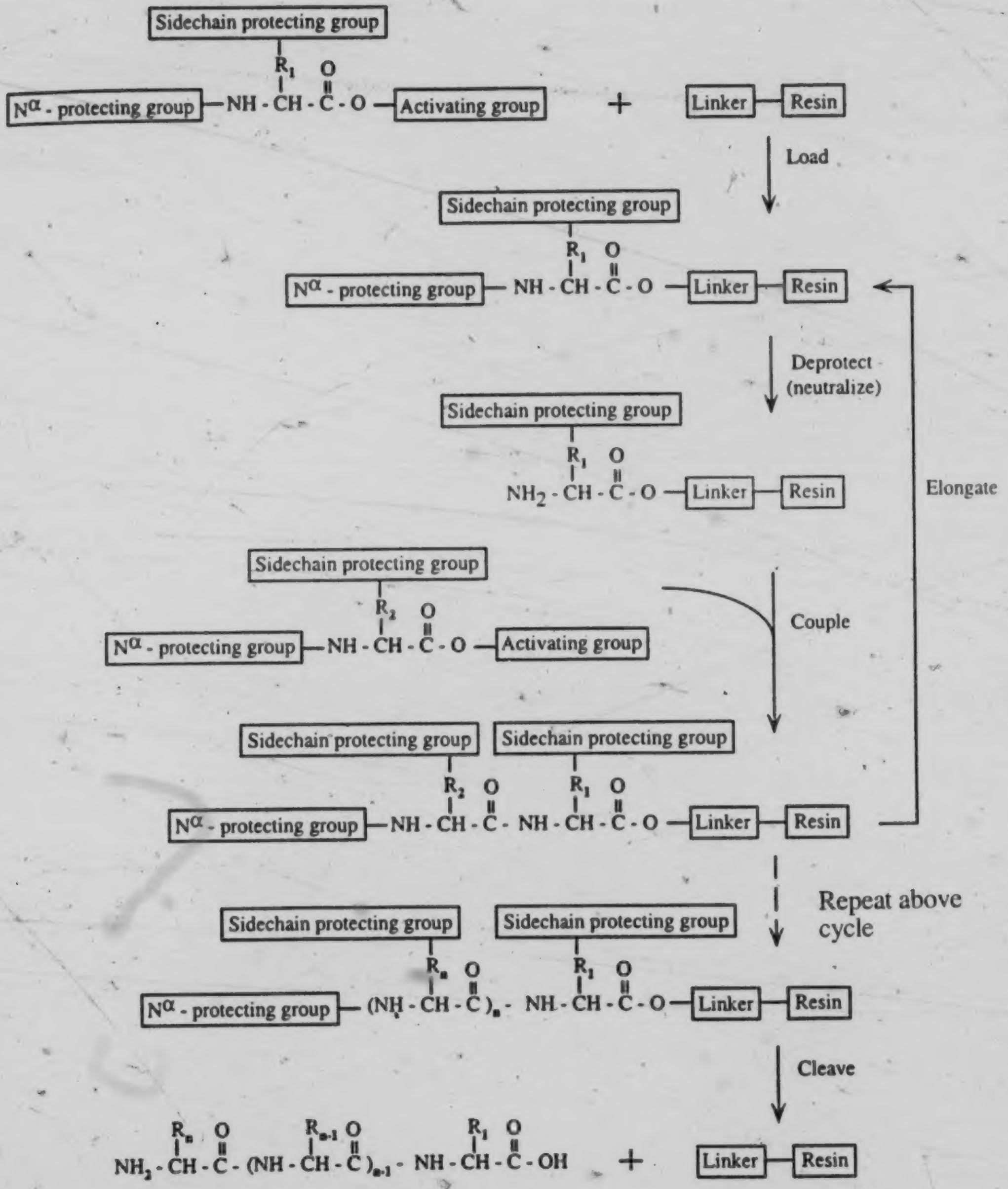
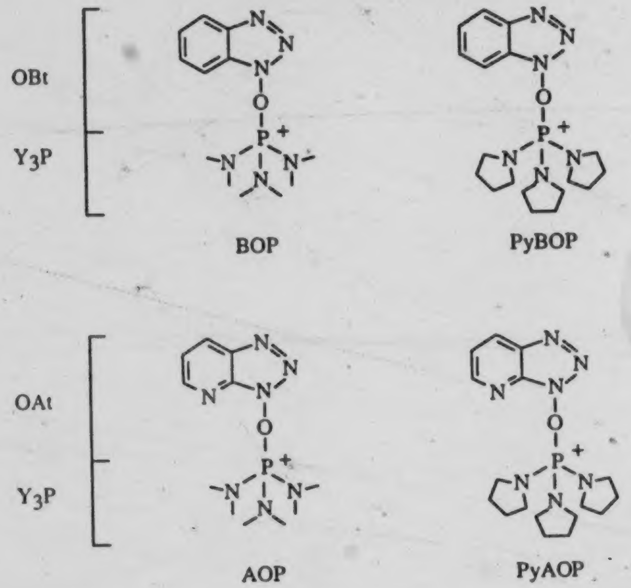

Sidechain protecting group
Nα - protecting group
NH - CH - C - O
Activating group
+
Linker
Resin
Load
Deprotect (neutralize)
NH2 - CH - C - O
Elongate
Couple
Repeat above cycle
Cleave
NH2 - CH - C - (NH - CH - C)n-1 - NH - CH - C - OH

Figure II-1. Schematic diagram depicting the various steps carried out during solid-phase peptide synthesis.

Structures of Phosphonium *In situ* Coupling Reagents


OBt
Y3P
BOP
PyBOP
OAt
Y3P
AOP
PyAOP

Structures of Aminium *In situ* Coupling Reagents

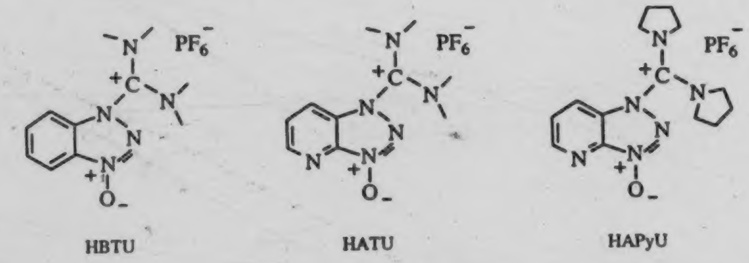

HBTU
HATU
HAPyU

Figure II-2. Structures of the *in situ* coupling agents used for peptide synthesis.

Formation of OBt and OAt Active Esters From Phosphonium *In situ* Coupling Reagents

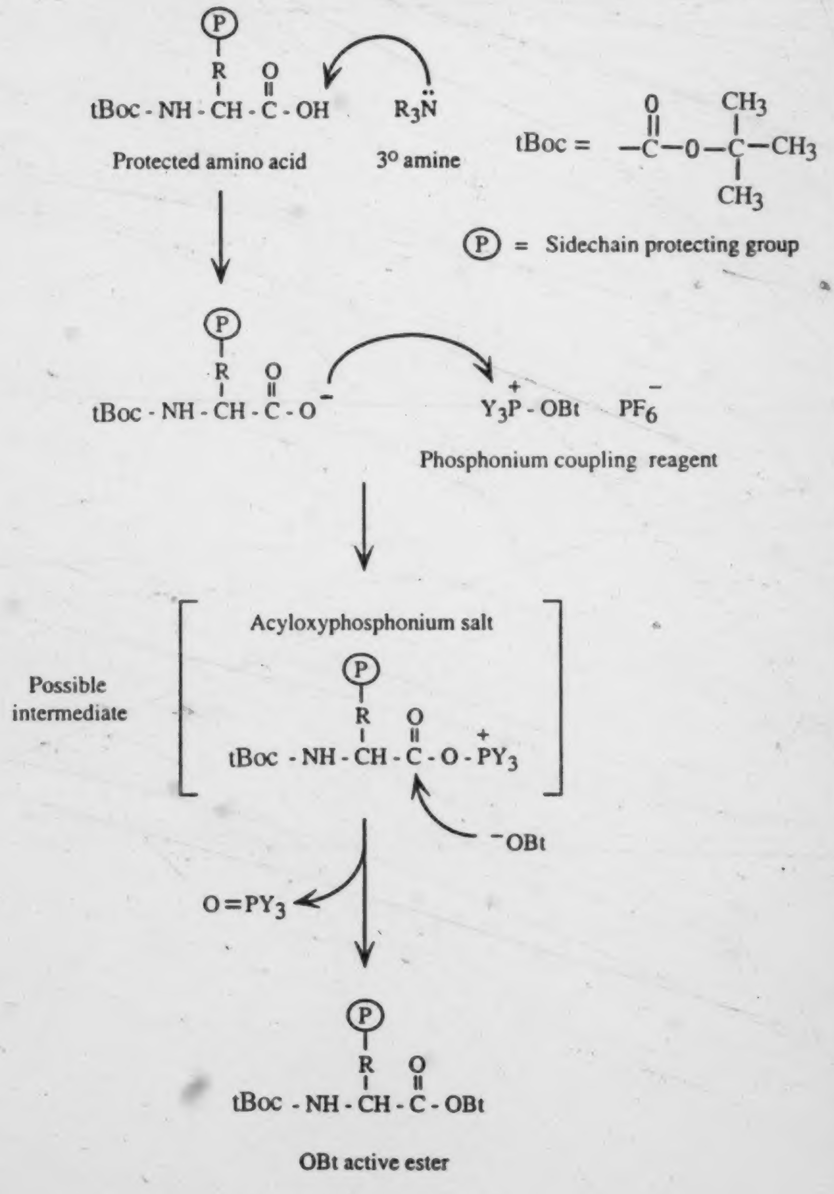

tBoc - NH - CH - C - OH
R₃N
Protected amino acid
3° amine
tBoc =
CH3
Ⓟ = Sidechain protecting group
tBoc - NH - CH - C - O⁻
Y₃P⁺ - OBt PF₆⁻
Phosphonium coupling reagent
Acyloxyphosphonium salt
Possible intermediate
tBoc - NH - CH - C - O - P⁺Y₃
⁻OBt
O=PY₃
tBoc - NH - CH - C - OBt
OBt active ester

Figure II-3. Activation of amino acids using *in situ* coupling reagents.

preparative C-8 column (250 x 10 mm I.D., 6.5-μm particle size, 300-Å pore size; SynChrom, Lafayette, IN) with a linear AB gradient (ranging from 0.2 to 1.0% B/min, depending on the peptide) at a flow rate of 2 ml/min, where solvent A is aqueous 0.05% TFA and solvent B is 0.05% TFA in acetonitrile. After lyophilization, half of the purified peptides containing oxidized methionines (Met(O)) were reduced using the procedure of Schechter (1986), and subsequently re-purified by RP-HPLC. Final peptide purity was verified by analytical RP-HPLC, amino acid analysis and electrospray quadrapole mass spectrometry.

Amino Acid Analysis

For amino acid analysis, purified peptides and proteins were hydrolyzed in 6 N HCl containing 0.1% phenol for 1.5 h at 160 °C in evacuated sealed tubes. Following HCl removal, samples were re-disssolved in appropriate volumes of sample dilution buffer (0.2 M NaCitrate/HCl, thiodiglycol 0.5%, benzoic acid 0.1%, pH 2 buffer) and run on a Beckman model 6300 amino acid analyzer utilizing post column ninhydrin detection analysis. Typical runs consisted of isocratic gradient steps using the following buffers: 0.2 M NaCitrate/HCl, pH 3; 0.2 M NaCitrate/HCl, pH 4.3; and 0.2 M NaCitrate, 0.9 M NaCl, 5% Phenol, pH 6 with coincident temperature changes between 46-66 °C.

Mass Spectrometry

The correct molecular masses of the purified peptides and proteins were confirmed by electrospray mass spectrometry using a Fisons VG Quattro mass spectrometer (VG Biotech, Altrincham, Cheshire, England) calibrated with horse heart myoglobin (Mr 16,951.5). Briefly, the samples (~10 μl) were injected and delivered to the analyzer using a 50 μl/min flow of 50:50 acetonitrile:water buffer in positive ion mode and scanned 7-10 times. Acquired data was then baseline subtracted, smoothed and centred. Determination of the transformed molecular weight was carried out by choosing "Manual mode processing" and picking adjacent ion peaks for the transformation. The molecular masses were determined with an uncertainty of +/- 2-3 Da.

Muscle Proteins

Rabbit skeletal muscle troponin (Tn) was purified according to the protocol of Ebashi *et al.*, (1971). Individual troponin subunits (TnT, TnI and TnC) were purified as described by Ingraham *et al.*, (1988). Rabbit cardiac αα-tropomyosin was purified as described by Pato *et al.*, (1981). G-actin was extracted and purified from rabbit skeletal muscle acetone powder as described by Spudich and Watts, (1971). Myosin subfragment 1 (S1) with associated alkali light chains (A1,A2) was prepared by the method of Weeds and Taylor (1975), as modified by Talbot and Hodges (1979).

Preparation of Peptides, Proteins and Complexes for ATPase Assays

Synthetic TnI peptides (1-2 mg/ml) were dissolved directly in 20 mM Tris, 6.5 mM KCl, 3.5 mM $MgCl_2$, 1 mM DTT, 0.01% sodium azide, pH 7.8 buffer (ATPase assay buffer). Actin (3.5 mg/ml), Tm (3 mg/ml), and TnC (3 mg/ml) (isolated as above) were dialyzed at 4 ^oC against several changes of the ATPase assay buffer prior to use. In the case of TnC, $CaCl_2$ was added into the dialysis buffer to make the final concentration 50 μM. TnI was initially prepared as described by Van Eyk and Hodges (1988) with final dialysis against ATPase assay buffer containing 100 mM KCl. Subsequent studies, however, showed that the protein could be dissolved directly into dH_2O (~5 mg/ml, pH ~ 4.5-5) and used without further preparation. Under the latter conditions, solubility of the TnI protein was markedly enhanced, oxidation of sulfhydryls was drastically reduced, and the overall activity did not differ from that observed from the first method. The binary complex between TnT and TnC (T•C complex) was prepared by dissolving TnT and TnC separately into ATPase (plus Ca^{2+}) buffer (20 mM Tris, 6.5 mM KCl, 3.5 mM $MgCl_2$, 1 mM DTT, 0.01% sodium azide, 50 μM $CaCl_2$, pH 7.8) containing 6 M urea. TnT and TnC concentrations were then determined by peak area integration at 210 nm using RP-HPLC (for more details see the actin-Tm centrifugation studies described below) and combined in a 1:2 molar ratio (instead of 1:1, see below), respectively. The T·C complex was subsequently dialyzed against the following series of ATPase (plus Ca^{2+}) buffers containing 1 M KCl; 0.5 M KCl; 0.25 M KCl; and finally 0.1 M KCl. The mole ratio of TnT to TnC of the final solution was 1:1.9. In our experience, it was found that a ~ two-fold molar excess of TnC to TnT was required in order to maintain solubility of TnT at the final salt condition required for this study. In all experiments, the amount of complex used was based upon the amount of TnC present in the complex in order to allow a direct comparison between Figures. III-3 and III-4 (i.e., neutralization by TnC in the presence or absence of TnT). Concentrations of all peptides and proteins were determined by amino acid analysis except for myosin S1, which was estimated by absorbance at 280 nm using $E_{1\%} = 7.9$ (Yagi *et al.*, 1967).

Acto-S1-Tm ATPase Assays

Release of inorganic phosphate (Pi) was quantitated using the colorimetric method developed by Heinonen and Lahti (1981), with the exception that Tris was used as the ATPase buffer and the pH of the reaction solution was increased from 7.0 to 7.8. The volume of all assays was 320 µl. The amount of actin, Tm and S1 used in each assay was: 1.5 nmol actin; 0.429 nmol Tm (dimeric) and 0.21 nmol S1 (giving a mole ratio of 7:2:1, respectively). The quantity of TnI peptide or protein, TnC or T·C complex varied according to the assay. For inhibition assays, increasing quantities of TnI or TnI peptides were added to the acto-S1-Tm solutions until maximum inhibition was reached. For the Ca^{2+}-dependent-TnC release assays, 0.858 nmol of TnI or TnI peptide (a mole ratio of 2:1 inhibitor:Tm) was added to the acto-S1-Tm solutions followed by the addition of increasing amounts of TnC. For the Ca^{2+}-dependent-T·C complex release assays, 0.858 nmol of TnI peptide or TnI protein was added to the acto-S1-Tm solutions, followed by the addition of increasing quantities of T·C complex (the amount of complex used was based upon the amount of TnC present in the complex). In both inhibition and release assays, appropriate aliquots of 100 mM KCl in ATPase buffer and 50 mM $CaCl_2$ in ATPase buffer were added into each assay solution to bring the final KCl concentration to 13 mM and the final $CaCl_2$ concentration to 0.4 mM (for release assays). Final assay solutions were pre-incubated at 25 °C for 10 min followed by initiation of the acto-S1-Tm reaction by the addition of 20 µl of 20 mM ATP in ATPase buffer solution. The reactions were allowed to proceed for 10-45 min depending on the assay, after which the reaction was stopped by the addition of 2.44 ml of a 2:1:1 (v/v/v) acetone: 5 N sulfuric acid : 10 mM ammonium molybdate solution and 240 µl of 1 M citric acid. The absorbance of individual tubes was read at 355 nm and the quantity of liberated inorganic phosphate determined. Endogenous S1 ATP hydrolysis was also determined for each assay time and subtracted from the data. Each ATPase data point was determined in triplicate. The acto-S1 and acto-S1-Tm ATPase activity's under the above conditions varied between 1.1-1.3 S^{-1} and 0.4 - 0.43 S^{-1} respectively.

Ca^{2+} Sensitivity Assays

The release of the inhibitory activity of the TnI peptides by TnC or T•C complex as a function of Ca^{2+} concentration was determined by adding 0.858 nmol of TnI or TnI peptide with 1.716 nmol of TnC or T•C complex to the acto-S1-Tm solution (as described above) to give a final solution composition of 7:2:1:4:8 (mole ratio of actin:Tm:S1:TnC or T•C:inhibitor). EGTA in ATPase assay buffer was added to make the final concentration 80 µM . Appropriate aliquots of $CaCl_2$ in ATPase buffer were then added to give free Ca^{2+} concentrations between pCa (-log [Ca2+]) values of 8 and 4.5, and the change in acto-S1-Tm activity determined (as above). pCa values for each calcium concentration were calculated using the program "comics" written by B. D. Sykes, University of Alberta, using binding constants similar to those described by Van Eyk *et al.* (1997). pCa_{50} values for each titration curve were determined by fitting of the data to a sigmoidal equation using the data fitting program "Kalidagraph" (Synergy Software, Reading, PA).

Rp_{1-40} Competition Assay

Competition between the N- and the C-terminal regions of TnI for TnC were carried out by adding 0.858 nmol of inhibitory peptide/protein with 1.716 nmol of TnC to actin-S1-Tm solutions (as described above) in the presence of 0.4 mM $CaCl_2$. Increasing amounts of Rp_{1-40} peptide were then added to the reaction mixtures, and the final solution incubated 10 min prior to initiation of the acto-S1-Tm activity with ATP (as described above). The ability of Rp_{1-40} to bind and displace the inhibitory peptides from TnC was monitored by the decrease in acto-S1-Tm ATPase activity.

Actin-Tm Centrifugation Studies

F-Actin, Tm and synthetic TnI peptides were prepared as described above, with the exception that the KCl and $MgCl_2$ concentrations in the ATPase assay buffer were increased to 100 mM and 5 mM, respectively as described by Van Eyk *et. al.* (1997). Each binding point was determined by mixing 5 nmol of actin, 0.429 nmol of Tm (dimeric) (a mole ratio of 7:2, respectively) and 0-30 nmol of synthetic TnI peptide in a total assay volume of 175 µl. The actin-Tm-peptide solutions were allowed to incubate at room temperature for 10 min prior to being spun for 30 min at 25 psi on a Beckman (San Ramon, California) airfuge using a 18 Å-100 rotor. The supernatant was removed and the pellets re-dissolved in 100 µl of 0.05% aqueous TFA. 80 µl of the "pellet" solution was then injected onto a Hewlett-Packard (Avondale, PA) 1090 HPLC pumping system equipped with a Zorbax SB 300 C-8 reversed-phase column (250 x 4.6 mm I.D., 6.5-µm particle size, 300-Å pore size Rockland Technologies, Wilmington, DE). The various proteins and peptides were eluted using a 2%B/min linear gradient from 0-70%B at a flow rate of 1 ml/min, where solvent A is aqueous 0.05% TFA and solvent B is 0.05% TFA in acetonitrile. Peak areas were determined by integration at 210 or 280 nm and converted to nmol of each protein or peptide using a standard curve obtained for each protein or peptide.

TnC-TnI Peptide Affinity Chromatography

Rabbit skeletal TnC was immobilized on cyanogen bromide-activated Sepharose 4B (Pharmacia, Uppsala, Sweden) according to the procedure suggested by the manufacturer. After washing to remove non-specifically bound protein, the coupled TnC-Sepharose was packed into a column (0.5 cm x 2.5 cm) and equilibrated with several column volumes of starting buffer containing 5 mM TrisHCl, 50 mM NaCl, 5 mM $CaCl_2$ or 5 mM $MgCl_2$, pH 7, at a flow rate of 0.5 ml/min. 200 μg of each of the TnI peptides dissolved in starting buffer (see Table 2 for peptides analyzed) were applied to the column followed by several column volumes of starting buffer. After obtaining a stable baseline, peptides retained on the column were eluted using a step gradient of starting buffer containing 1 M NaCl (see footnotes to Table 2 for more details). Peptides which were not eluted under these conditions were removed by washing with an aqueous solution of 5 mM TrisHCl, pH 7.8, containing 6 M GdnHCl and 10 mM EDTA. Elution time and conditions for each peptide were determined by monitoring the column effluent at 210 nm using a Hewlett-Packard 1090 diode array detector. Peptides which required GdnHCl to be eluted were verified by analytical RP-HPLC (see binding studies for conditions).

Circular Dichroism Measurements.

Circular dichroism (CD) spectra were recorded on a Jasco J-720 spectropolarimeter (Jasco Inc., Easton, MD) interfaced to an Epson Equity 386/25 computer running the Jasco DP-500/PS2 system software (ver. 1.33a). The temperature-controlled cuvette holder was maintained at 20 °C (unless otherwise stated) with a Lauda model RMS circulating water bath (Lauda, Westbury, NY). The instrument was routinely calibrated with an aqueous solution of re-crystallized *d*-10-(+)-camphorsulphonic acid at 290.5 nm. Results are expressed as mean residue molar ellipticity $[\Theta]$ ($\mathrm{deg \cdot cm^2 \cdot dmol^{-1}}$) calculated from the equation:

$$[\Theta]=(\Theta_{obs} \times \mathrm{MRW})/(10 \times l \times c)$$

where Θ_{obs} is the observed ellipticity expressed in millidegrees, MRW is the mean residue molecular weight (molecular weight of the peptide divided by the number of amino acids), l is the optical path length in cm, and c is the final peptide concentration in mg/ml.

For wave scans, data was collected from 190 to 255 nm at 0.05 nm intervals, and the average of 10 scans reported. For wave scans in the prescence of 50% trifluoroethanol (TFE), samples were diluted 1:1 (v/v) with neat TFE.

Urea and GdnHCl denaturation studies were carried out by monitoring the ellipticity at 222 nm as a function of denaturant concentration. Each ellipticity reading involved the averaging of 10 scans. Samples were prepared by preparing mixtures of a stock solution of peptide in buffer (0.1 M KCl, 0.05 M PO_4, +/- 0.002 M DTT, pH 7), buffer alone, and a solution of 8 M GdnHCl (or urea) in buffer and mixing various ratios of buffer and denaturant solutions to give appropriate final denaturant concentrations.

Temperature denaturation studies were carried out similarly by monitoring the ellipticity at 222 nm but as a function of temperature. Typical experiments involved measurements from 4 °C to 85 °C.

Concentration dependence studies were carried out by monitoring the change in α-helical content at 222 nm at various peptide concentrations. Typically peptide concentrations ranged bewteen 10 μM and 1 mM. Correspondingly, pH titrations followed by CD spectroscopy were done by making up several buffer solutions at the desired pH values using combinations of K_2HPO_4 and KH_2PO_4 (50 mM total) and adjusted with 6 N HCl. Then 10 μl of the peptide stock solution (~5 mg/ml) was added to 90 μl of each buffer solutions. The final pH of each analyzed solution by CD was verified using a pH meter equiped with a micro-probe.

In general, the cell path length used was 0.05 cm for the data points acquired for denaturation and pH/concentration dependance studies monitored at 222 nm, and 0.02 cm for CD spectral scans. All peptide concentrations were determined by amino acid analysis on a Beckman model 6300 amino acid analyzer as described above.

Calculation of Denaturant Midpoints, Free Energy of Unfolding ($\Delta G_u^{H_2O}$) and Differences in the Free Energies of Unfolding ($\Delta\Delta G_u$).

Denaturation midpoint values for the unfolding of the various peptides from folded (f) to unfolded (u) states was determined by following the change in molar ellipticity at 222 nm at 20 °C. Ellipticity readings were normalized to the fraction of the peptide folded (f_f) or unfolded (f_u) using the standard equations:

$$f_f = ([\theta] - [\theta]_u)/([\theta]_n - [\theta]_u)$$

$$f_u = (1 - f_f)$$

where $[\theta]_n$ and $[\theta]_u$ represent the ellipticity values for the fully folded and fully unfolded species, respectively. $[\theta]$ is the observed ellipticity at 222 nm at any denaturant concentration. Each curve was then fit to a sigmoidal equation (e.g. (max-min)/(1+(denaturant conc./midpoint)^slope)+min).

Calculation of the $\Delta G_u^{H_2O}$ (the free energy of unfolding in the absence of guanidine hydrochloride) was estimated by extrapolating the free energy of unfolding at each denaturant concentration to zero concentration assuming they are linearly related by the equation $\Delta G_u = \Delta G^{H_2O} - m[\mathrm{GdnHCl}]$ (Pace, 1986; Shortle, 1989). ΔG_u was calculated from the equation: $\Delta G_u = -RT \ln(K_u)$

where $K_u = (f_u/f_f)$ for the oxidized peptides, and $(2P_t(f_u^2/(1-f_u))$ for reduced peptides (De Francesco *et al.*, 1991) and P_t is the total peptide concentration (M).

When the difference in free energy change ($\Delta\Delta G_u^{H_2O}$) was sought between a series of analogs relative to a specific analog, the $\Delta\Delta G_u^{H_2O}$ values were calculated using the equation given by Serrano and Fersht (1989):

$$\Delta\Delta G_u = ([\text{denaturant}]_{1/2,A} - [\text{denaturant}]_{1/2,B})(m_A + m_B)/2$$

where $[\text{denaturant}]_{1/2,A}$ and $[\text{denaturant}]_{1/2,B}$ are the midpoints of unfolding of analog A and B respectively derived from the plots of fraction folded versus GdnHCl concentration, and m_A is the slope determined from the equation $\Delta G_u = \Delta G_u^{H_2 0} - m$ [denaturant] (as described above). This approach calculates the value of $\Delta\Delta G_u$ at the denaturant concentration half-way between the $[\text{denaturant}]_{1/2}$ values of the peptides being compared. This method has been shown to be more accurate for determining small differences in protein stability because small errors from extended extrapolation are avoided and linearity between ΔG_u and denaturant concentration is assumed across only a small range of denaturant concentration.

Size Exclusion Chromatography with Laser Light Scattering

Molecular weights of the peptides in aqueous solution were determined by size-exclusion chromatography (SEC) with laser light scattering. SEC was carried out on a Superose 12 column (1.0 cm x 30.0 cm) from Pharmacia at a flow rate of 0.5 ml/min. at room temperature. The eluent was a 100 mM KCl, 50 mM K_2HPO_4, pH 7 buffer. The effluent from the column was monitored using either a Hewlett Packard UV-visible spectrophotometer at 210 nm, or a Dawn F multiangle laser light scattering photometer connected in series with a Optilab 903 refractometer. Determination of molecular weights (by laser light scattering) was carried out according to the methodology described by Farrow *et al.* (1994).

Preparation of Oxidatized peptides.

Formation of homo-two-stranded disulfide bridged molecules was carried out by dissolving 5 mg of each peptide into 2 ml of 100 mM NH_4HCO_3 pH 8 buffer and the reaction vessel stirred overnight at room temperature. Oxidized peptides were re-purified by RP-HPLC and oxidation verified by mass spectrometry (as described above).

Helical Propensity / Hydrophobicity Analysis

The score for α–helical propensity and hydrophobicity (occurring in a 3-4 repeating pattern) was calculated for residues 280-420 of the kinesin protein. Each α–helical propensity data point was obtained by an iterative process involving the summing of 11 individual α–helical propensity scores (Monera *et al.*, 1995) for the sequence starting at residue position 280. Subsequent data points were then obtained by shifting 1 residue toward the COOH-terminus and repeating the process. Hydrophobicity values, occurring in a 3-4 repeating pattern, were calculated in a similar manner by summing the hydrophobicity scores (Monera *et al.*, 1995) for the first 6 residues that occur in a "3-4" repeating pattern (i.e. residues 1,4 8,11, 15 and 18) . Subsequent data points for the same face were then calculated by shifting first 3 residues (4, 8, 11, 15, 18, and 22) then 4 residues (8, 11, 15, 18, 22, and 25) and so forth throughout the sequence in order to maintain the same reading frame. Calculation of the other six faces was carried out by starting +1 residue toward the COOH-terminus (2, 5, 9, 12, 16 and 19) *etc.* to obtain data for the second face and another +1 residue shift (3, 6, 10, 13, 17 and 20) *etc.* for face three. This process is repeated to obtain data for faces 4 to 7.

Sedimentation Equilibrium and Velocity Studies .

Conventional sedimentation equilibrium experiments were performed on a Beckman Model XL-I analytical ultracentrifuge (Beckman, Palo Alto) at 20°C. Standard charcoal-filled Epon double sector centerpieces (12 mm) were used, and a sample volume of 120 µL was added into each sample sector. Peptide samples were initially dialyzed against the desired buffer, at 4°C overnight prior to analysis. Experiments were conducted at various rotor speeds depending on the size of the complex (~18,000 to 32,000 rpm). Data were acquired by averaging 5 radial scans at a spacing of 0.001 cm. Equilibrium was confirmed by the superimposition of the concentration gradient measured after 20 and 24 h. The data were fit to the following equation:

$$c_r = c_0 \exp\{[M_{w,app}(1- \bar{v}_2\rho)(r^2-r_0^2)\omega^2]/2RT\}$$

where $M_{w,app}$ is the apparent molecular mass, $\bar{v}_2$ is the partial specific volume, c is molar concentration, r is radial distance, and ρ is the solvent density. The partial specific volume of each sample was calculated from amino acid composition using the program SEDINTERP (Hayes *et. al.* 1998).

Sedimentation velocity experiments were carried out at 20 °C and 48K rpm using a Beckman XL-I analytical ultracentrifuge and interference optics as described in the operating manual (Beckman Instruments, 1997a). Due to the low sedimentation coefficients of the samples, 2-sector charcoal-filled Epon synthetic boundary centerpieces were employed to layer dialysate onto the protein solution to provide a depleted meniscus. Runs were performed for 4 h during which a minimum of 20 scans were recorded. The sedimentation velocity data was analyzed using the Second Moment Method contained in the Beckman Analysis Program (Beckman Instruments, 1997b) to determine the sedimentation coefficient from the following relationship:

$$s\varpi^2 = \frac{d\ln r}{dt}$$

where s = observed sedimentation coefficient, ϖ = angular velocity, r = radial position of the solvent/protein boundary and t = the time in seconds. By measuring the movement of the boundary over time and plotting ln r versus $\varpi^2 t$, a straight line should be obtained with slope = s.

The program SEDINTERP (Hayes *et al.*, 1998) was then used to calculate $S^{o}_{20,w}$, f/f_0 and a/b according to the following relationships:

$$S_{20,w} = s\left(\frac{\eta_t}{\eta_{20}}\right)\left(\frac{\eta_{sol}}{\eta_w}\right)\left(\frac{1-\bar{v}\rho_{20,w}}{1-\bar{v}\rho_{t,sol}}\right)$$

where $S_{20,w}$ is the sedimentation coefficient corrected to water at 20 °C, s = the observed sedimentation coefficient, η_t = the viscosity of water at the run temperature, η_{20} = the viscosity of water at 20 °C, η_{sol} = the viscosity of the sample solution at temperature t', η_w = the viscosity of water at temperature t', $\bar{v}$ = the partial specific volume calculated using the method of Cohn and Edsall (1943), $\rho_{20,w}$ = the density of water at 20 °C and $\rho_{t,sol}$ = the density of the sample at the run temperature.

$$S^{\circ}_{20,w} = \frac{S_{20,w}}{(1-kC_p)}$$

where $S^{o}_{20,w}$ is $S_{20,w}$ extrapolated to infinite dilution, k = an empirical coefficient (0.009 ml/mg for spherical; proteins) and C_p = the protein concentration in mg/ml.

$$f = \frac{M(1-\bar{v}\rho)}{Ns}$$

where f = the translational frictional coefficient, N = Avagadro's number, M = the molecular weight and s = the sedimentation coefficient.

$$f_0 = 6\pi\left(\frac{3M\bar{v}(1-\delta)}{4\pi N}\right)^{1/3}$$

where f_0 = the calculated frictional coefficient of a sphere having the same volume as the protein molecule and δ = the degree of hydration (g H_2O/ g protein). SEDINTERP calculates the degree of hydration from the amino acid composition of the protein using the method of Kuntz (1971). The ratio f/f_0 provides a measure of the degree of asymmetry of the molecule. Modeling the particle as a prolate ellipsoid of revolution allows estimation of the axial ratio (a/b) using a power series approximation of the tabulated data for a/b as a function of f/f_0 (Cantor & Schimmel, 1980).

High Performance Size Exclusion Chromatography of the Model Coiled-Coil Peptides:.

Determination of the oligomeric state(s) of each peptide analog in benign buffer conditions (50 mM potassium phosphate, 100 mM KCl, pH 7) as well as in the presence of various concentrations of GdnHCl denaturant in buffer was carried out using a Superdex™75 HR 10/30 (10 mm x 300 mm bed size) high performance size-exclusion column (Pharmacia, Uppsala, Sweden) and a Hewlett Packard 1090 HPLC work station.

The column was conditioned in several column volumes at a flow rate of 0.4 ml/min prior to injection. Each peptide was either injected from the stock solution (~1mM) or 10 fold diluted (~0.1 mM) in running buffer and the chromatogram collected at 215 nm. The peak area indicated a further 10 fold dilution occurred within the period of the run time. To correct for bed volume changes due to changes in the ionic strength of the various buffers used, two random coil internal peptides standards were also co-injected with each peptide analog and the chromatogram corrected to the elution times of the standards: S1= Ac-(GLGLG)$_6$-amide (30 residues, MW 2534 Da); S2= Ac-KYGLGGAGGLK-amide (11 residues, MW 1062 Da). Molecular weights were estimated based on their correlation in retention time to several in house coiled-coil standards of various sizes.

RP-HPLC Analysis of the Model 'a' and 'd' Peptide Analogs

Determination of the relative hydrophobicity of each peptide analog was performed by reversed-phase HPLC (RP-HPLC) on a Zobax Eclipse XDB-C8 column (150 x 4.6 mm I.D., 5-μm particle size, 300-Å pore size); Hewlett-Packard, Little Falls Site, DE, USA) using a Hewlett-Packard 1090 HPLC work station. Briefly, 5 μg of each carboxamidomethyl modified peptide (see below) was run individually as well as together in appropriate groups to determine the retention time of each analog relative to that of the Ala substituted analog. Each run consisted of a linear AB gradient of 1.0% B/min. at a flow rate of 1 ml/min and 70 °C, where solvent A is aqueous 50 mM KH_2PO_4, 100 mM $NaClO_4$, pH 7 and solvent B is 50 mM KH_2PO_4, 100 mM $NaClO_4$, pH 7 in 50% acetonitrile.

In order to derive retention times independent of the stability of the coiled-coils, each oxidized peptide was first reduced with 2 equivalents of dithiothreitol (DTT) and then capped (i.e. the sulfhydryl group of the cysteine side chain) with 4 equivalents of iodoacetamide to prevent re-oxidation. Excess DTT was then added to neutralize any remaining iodoacetamide.

Coiled-coil Prediction Analysis

Prediction of the location and stability of the coiled-coil regions within the kinesin neck regions was carried out as that described by Tripet *et al.*, (1999). The predictive method is based on the ability of sequence regions to score above pre-defined cut-off threshold values using various sized window widths (e.g 21, 28, 35 and 42). Scores are tabulated iteratively using helical propensity values (Monera *et al.*, 1995) for positions b,c,e,f,g and hydrophobic core stability values (Wagschal *et al.*, 1999 and Tripet *et al.*, 1999) for positions "a" and "d" of the heptad repeat (denoted abcdefg). An ~ 2.5 fold weighting was given to positions "a" and "d" over that of the other positions. Plots are generated by plotting the cummulative score (of the choosen window size) originating from that residue versus the residue position. The final predictied stability profiles are based on a 7 residue (one heptad) window width for the predicted region.

Recombinant Engineering (enzymes, oligomers and general techniques)

Restriction enzymes and DNA modifying enzymes were purchased from GIBCO BRL (Life Technologies, Grand Island, NY). Deoxyribonucleotides were prepared using an Applied Biosystems (Foster City, CA.) automated DNA synthesizer and isolation of plasmid DNA and routine manipulations were carried out according to standard procedures (Maniatis *et al.*, 1982) Dideoxy-sequencing (Maxam and Gilbert, 1977) of the cloned plasmid constructs were achieved using a modified PCR cycle sequencing protocol (Craxton, 1991) employing biotinylated primers and chemiluminescent detection (Amersham, Oakville, Ontario, Canada).

Bacterial Stains and Plasmids

E.coli K-12 expression strain JM83 [*ara*, Δ(*lac-pro AB*),*rpsL* (=*str A*) (Φ80 *lacZ*ΔM15)] (Yanisch-Perron *et al.*, 1985) was obtained from New England Biolabs (Mississauga, Ontario, Canada) and used as the host cell in all prokaryotic expression experiments. Cells were cultured in Luria broth (LB) media (Sambrook *et al.*, 1989) supplemented with carbenicillin (100μg/ml).

Plasmid pASK40 (generously provided by Dr. Arne Skerra, Max-Planck-Institut für Biophysik, Germany) was used for plasmid construction.

Synthetic E-coil Gene

A DNA sequence (Table IX-I) was constructed from back translation of the *de novo* designed E-coil protein sequence (Table IX-II) using a codon bias based on the highly expressed *Pichia pastoris* genes (Koutz *et al.*, 1989). To avoid the creation of an excess number of restriction sites, four overlapping oligonucleotides were synthesized and purified by electroelution on a 12% polyacrylamide gel containing 7 M urea. Oligonucleotides #1 and #2 (Table IX-I) were annealed together by their complementary sequences and the remaining complementary strands synthesized by klenow fragment. An identical protocol was carried out for oligonucleotides #3 and #4. The double stranded fragments were digested with restriction enzyme Alw21 I to create sticky ends at one end of each fragment, and the two fragments were subsequently ligated with T4 DNA ligase to give a complete gene of 177 base pairs that could be detected on a 4% agarose gel. The gene was excised and electroeluted from the gel slice, then digested simultaneously with EcoR1 and BamH1 to produce a 166 bp fragment suitable for cloning into the plasmid vector.

Plasmid Construction

The plasmid vector pRLDE was constructed by the following steps: 1) the plasmid pASK40 was digested with EcoR1 and the 5' ends blunted with Mung-Bean Nuclease. The blunted plasmid was then digested with Hind III and religated with the synthetic 43 bp annealed deoxyribonucleotide fragment (Table IX-1 oligonucleotides #5 and #6), creating pRLD. 2) pRLD was digested with EcoR1 and Bgl 1 and then force cloned with the synthetic E-coil gene (see above) to produce the *E. coli* E-coil vector pRLDE (Figure IX-3).

Transfection and Screening

The ligated plasmids were transformed into *E. coli* by heat-shock treatment (Yanisch-Perron *et al.*, 1985) and plated on selective LB plates supplemented with carbenicillin (100 µg/ml). Insertion of the E-coil gene was verified by restriction digest analysis and subsequent DNA sequencing of purified plasmid DNA of successful transformants.

Construction of the PAK-pilin Gene Fusion

Two complementary oligonucleotides (oligonucleotides # 9 and #10, Table IX-1) encoding the 17 amino acid pilus epithelial binding domain of the *Pseudomonas aeruginosa* strain K pilin (PAK-pilin)[residues 128-144,(Lee *et al.*, 1994)] were synthesized. Equal molar ratios of oligonucleotides dissolved in sterile distilled H_20 were mixed together, heated at 80°C for 5 min and finally allowed to anneal by slow cooling to room temperature. The annealed oligonucleotides were electrophoresed on a 4% agarose gel and a 54 bp fragment was excised and eluted from the gel slice by a modified freeze squeeze method (Maniatis *et al.*, 1982). Briefly, the excised band was placed in a spin filter column (BioRad, Mississauga, Ontario, Canada) and set on liquid nitrogen for 5 min. The filter was then placed into an eppendorf tube and spun at top speed in a microfuge. After ethanol precipitation, the eluate was treated with T4 polynucleotide kinase at 37°C for 30 min and the enzyme inactivated by heating at 65°C for 15 min. The phosphorylated fragment was then cloned into an EcoR1 digested, shrimp alkaline phosphatase treated pRLDE and subsequently transformed into *E.coli* to produce pRLDE-PAK-pilin clones on LB/carbenicillin plates. Orientation and sequence of the insert was verified by restriction analysis and sequencing of miniprepped DNA plasmid. Expression and purification of the PAK-pilin-E-coil peptide was carried out as described for the E-coil peptide.

Construction of the PAK-pilin(met) Gene Fusion

Two complementary oligonucleotides (oligonucleotides # 7 and #8, Table IX-1) encoding the 17 amino acid pilus epithelial binding domain (as described above) plus two adjacent methionine residues were synthesized. The two oligonucleotides were then treated in an identical fashion to that described above for the preparation of PAK-pilin gene fusion. The phosphorylated fragment was cloned into EcoR1 digested, shrimp alkaline phosphatase treated pRLDE and transformed into *E.coli* to produce pRLDE-PAK-pilin(met) clones on LB/carbenicillin plates. Orientation and sequence of the insert was verified by restriction analysis and sequencing of miniprepped DNA plasmid. Expression and purification of the PAK-pilin(met)-E-coil peptide was carried out as described for the E-coil and PAK-pilin-E-coil peptides.

Growth, Induction and Lysis of Bacteria Containing Recombinant Plasmids

E. coli cells harboring pRLDE plasmids were grown at 25°C, with shaking, in LB medium containing carbenicillin (100μg/ml) to an A_{550} of 0.5-0.6. Production of recombinant protein was then induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 1mM. The bacteria were then grown for an additional 3 h at 25°C with shaking. The expressed protein was obtained by a modified osmotic shock treatment (Ausubel *et al.*, 1992). Cells were centrifuged at 4,000 g for 10 min at room temperature and re-suspended in TES buffer (100mM TrisHCl, 5mM EDTA, 20% sucrose, pH 8.0) to a final volume of 80 ml per gram wet weight. Cells were shaken at 200 rpm at room temperature for 10 min. The suspension was then re-centrifuged and the pellet re-suspended in 5 mM ice-cold $MgSO_4$ (80 ml per gram wet weight). Cells were shaken for 30 min on ice and subsequently centrifuged at 8,000 g and 4°C for 10 min. The supernatant constituting the periplasmic fraction was then further clarified by passage through a 0.22-um filter and subsequently purified using the K-coil affinity column.

Preparation of the Biotinylated K-coil

5 mg (1.18 μmol) of the synthetically prepared K-coil peptide was dissolved in 5 ml of 50 mM KH_2PO_4, pH 7.0, buffer. To this solution was added 1.3 mg (2.36 μmol) of Biotin-BMCC (1-Biotin amido-4-[4'-(maleimidomethyl)cyclohexanecarboxamido]butane) (Pierce, Rockford, IL) that had been dissolved in 1 ml of 6 M GdnHCl in 50 mM KH_2PO_4 buffer, pH 7.0. The solution was stirred for 1h, after which the modified peptide was purified by RP-HPLC (see above).

Preparation of the Peptide affinity Matrix

1 g (110 μmol) of aminopropyl control pore glass resin (CPG)(125-μm particle size, 500-Å pore size, Sigma Chemical Company, St Louis, MO) was washed twice sequentially with 15 ml of dimethylformamide, 15 ml of dichloromethane and 15 ml of dimethylformamide. 15 ml of a 5% diisopropylethylamine/dichloromethane solution (v/v) was then added and the glass beads neutralized for 5 min and then drained. To 138 mg (1 mmol) of bromoacetic acid (Aldrich Chemical Company, Milwakaukee, WI) was added 2 ml of a 0.5 M dicyclohexylcarbodiimide/dichloromethane solution and the resulting solution stirred for 15 min. The solution was then filtered to remove precipitated dicyclohexylurea, and the filtrate diluted to 10 ml with dimethylformamide. The diluted solution was then added to the neutralized glass resin beads and the complete mixture stirred for 35 min, after which the solution was removed and the resin beads washed three times each with 15 ml of dichloromethane, dimethylformamide, methanol and finally, dichloromethane. A qualitative ninhydrin test was performed to verify coupling. Capping

of free silanol groups on the resin beads was carried out by adding 30 ml of a 75% pyridine/dichloromethane solution (v/v) to the bromoacetylated glass resin beads and then adding slowly 3 ml of chlorotrimethylsilane (Aldrich). The reaction was allowed to stir for 1 h after which the beads were washed three times each with 15 ml of dichloromethane, dimethylformamide, methanol and finally, dichloromethane.

100 mg (23.7 μmol) of the synthetically prepared K-coil peptide was dissolved in 30 ml of an aqueous 50 mM NH_4HCO_3 solution, pH 8.0, and added to 1 g of the activated glass resin beads. The peptide resin solution was then stirred for 30 min, after which the resin was drained and washed with 15 ml of distilled H_2O (dH_2O). Remaining unreacted bromoacetyl groups were capped by adding 20 ml of a 1% solution of β-mercaptoethanol (v/v) in 50 mM NH_4HCO_3, pH 8.0, to the glass resin beads, and the mixture stirred for another 15 min. The resin was drained and washed five times each with dH_2O, dH_2O containing 0.1% TFA, ethanol and finally, acetone. The final K-coil affinity chromatography resin was air dried and a small sample of the resin removed for amino acid analysis. Final resin substitution was calculated to be 12.5 μmol of peptide/g resin.

Preparation of the Affinity Column

Two columns were prepared for affinity purification chromatography: (1) a stainless steel RP-HPLC guard column (20 mm x 5 mm I.D.) was used for small scale analytical runs; (2) a stainless steel RP-HPLC column (200 mm x 4.6 mm I.D.) was used for large scale purification of 1 L or greater volumes of culture. Each column was dry-packed with the prepared K-coil affinity chromatography resin (50 mg and 1 g, respectively) and then washed at a flow rate of 0.5 ml/min with the following solutions: dH_2O, 5.0 M GdnHCl in 50 mM KH_2PO_4, pH 7.0, 50% acetonitrile/H_2O containing 0.1% TFA(v/v/v), and, finally, dH_2O. The mobile phase used for column storage was 50% acetonitrile/H_2O (v/v).

Purification of Recombinant Peptides

Affinity purification of the recombinantly expressed peptides was carried out using a Varian (Varian, Walnut Creek, CA) Vista Series 5000 liquid chromatograph. The affinity columns were first conditioned in a several column volumes of starting buffer of 10 mM KH_2PO_4, pH 6.0. Filtered (through a 0.22 µm filter) periplasmic extract obtained from osmotically shocked cells or whole cell lysate from french pressed cells were loaded onto the affinity column at a flow rate of 0.2-0.5 ml/min . After loading, the columns were first washed with 0.5 M KCl in 10 mM KH_2PO_4, pH 6.0, followed by 80% acetonitrile in 10 mM KH_2PO4, pH 6. Final elution of bound material was carried out by washing the column with 50% aqueous acetonitrile containing 0.1% TFA (v/v/v). Volume sizes for each wash step and final elution were 5 ml for the small affinity column and 10 ml for the large column at a flow rate for both of 0.5 ml/min.

RP-HPLC Analysis of the Affinity Purification Method

Crude periplasmic extract (10 ml) of the recombinantly expressed E-coil fusion tag was loaded onto the analytical affinity chromatography column at a flow rate of 0.2 ml/min. After loading, the column was washed with each of the various wash solutions (see above) and the eluent from each step collected for RP-HPLC analysis. The sample loadings for RP-HPLC analysis were as follows: 1 ml of the crude periplasmic extract, 1 ml of the crude periplasmic extract after passage through the affinity column (breakthrough), 2.5 ml of the 5 ml KCl wash, 1 ml of the 80% aqueous acetonitrile wash, 1 ml of the first elution wash, 1 ml of the second elution wash, and 1 ml of the first elution wash after re-running the breakthrough. Note that in the case of the 80% and 50% aqueous acetonitrile washes, the eluates were first diluted two-fold with dH_20 and lyophilized. The samples were then re-dissolved prior to injection in 2 ml of aqueous 0.1% TFA.

Analysis of the recombinantly expressed PAK-pilin-E-coil fusion peptide was carried out in a similar manner with one exception. 5 ml of crude periplasmic extract and 5.0 ml of the crude periplasmic extract after passage through the affinity column (breakthrough) were injected for RP-HPLC analysis instead of 1 ml.

RP-HPLC analysis of the eluates was carried out on a Hewlett-Packard (Avondale, PA) 1090 HPLC pumping system equipped with a Zorbax SB 300 C-8 column (250 x 4.6 mm I.D., 6.5-μm particle size, 300-Å pore size Rockland Technologies, Wilmington, DE). Each run was performed with a flow rate of 1.0 ml/min and a 1% B/min gradient from 0-60%, where solvent A is 0.1% TFA in H_2O and solvent B is 0.1% TFA in acetonitrile. Absorbance data were collected at 210 nm and each run plotted with a vertical axis of 1000 AU.

Purification from Whole-Cell Lysate

Two 200 ml cultures of *E.coli* cells containing either the E-coil fusion tag vector or the PAK-pilin-E-coil fusion peptide vector were grown to an A_{550} of 0.5, induced for 4h (as described previously), and then concentrated by centrifugation at 4000 g for 10 min. The cell pellet was resuspended in 3 ml of a 100 mM Tris-HCl, 5 mM EDTA, 100 ug PMSF, pH 8.0 buffer and french pressed to liberate all cellular proteins and components. The complete cellular lysate was then clarified by centifiugation in a microfuge at top speed for 5 min and the soluble fraction applied to the K-coil affinity column for purification (see above). RP-HPLC analysis of the final elution washes is described in the figure legend of Figure IX-6.

SDS-PAGE

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed with 15% acrylamide gels in a mini-gel apparatus (Mini-protein II Dual Slab Cell from Bio-Rad) as described by Laemmli (Laemmli, 1970). Crude periplasmic extract from *E coli* JM83 cells containing expressing E-coil and PAK-pilin-E-coil peptides as well as their K-coil affinity purified fractions were fractionated by SDS-PAGE at a constant voltage of 200V. Gels were stained with Coomassie Blue (R-250, Bio-Rad) for visualization .

Western Blotting using Anti-PAK-pilin

Proteins on the SDS-PAGE gel were transferred to nitrocellulose membrane using the protocol of Towbin *et al.* (1979) with a Mini Trans-Blot Electrophoretic Transfer Cell (Bio-Rad). The transfer was completed after 30 min under a constant current of 300 mA (Model 200/2.0 Power supply from BioRad). Excess binding sites on the membrane were blocked by incubation of the blots with a blocking solution consisting of 50 mM Tris-HCl, pH 7.5, containing 150 mM NaCl, 0.05% (v/v) Nonidet-P40, 0.25% (w/v) gelatin and 3% (w/v) BSA at room temperature for 1 h in an incubator shaker (model G25 Gyroshaker, New Brunswick Scientific, NJ) set at 100 rpm. The membrane was washed twice at room temperature with 10 mM Tris-HCl buffer, pH 7.5, containing 0.1% Tween-20 and 0.05% (w/v) BSA (TTBS). Monoclonal antibodies, PK99H raised against the *Pseudomonas aeruginosa* strain K pilin (PAK pilin), were diluted with TTBS (1:500) and added to the blots and incubated for 1 h at room temperature with gentle shaking. The blots were washed three times with TTBS. A goat anti-mouse IgG(H+L)-alkaline phosphatase conjugate (Jackson ImmunoResearch Laboratory, Inc., West Grove, PA) diluted 1:10,000 with TTBS, was incubated as described above. The strips were washed 3 times with TTBS followed by a final wash with Tris-buffered saline. Antibody binding was visualized by the addition of alkaline phosphatase substrates (nitro blue tetrazolium chloride and 5-bromo-4-chloro-3-indolyl phosphate dissolved in 100 mM Tris-HCl, pH 9.5, containing 100 mM NaCl and 5 mM $MgCl_2$). Color development was terminated by rinsing the nitrocellulose strips with de-ionized water.

Western Blotting using Biotinylated K-coil

Extracted and purified peptides run by SDS-PAGE were transferred to nitrocellulose membrane as described above. The proteins on the membrane were denatured by soaking the membrane for 1 h in PBS-ED buffer (20 mM potassium phosphate, 200 mM KCl, 1 mM EDTA, 1 mM DTT, pH 7.9) containing 6 M guanidine HCl. The proteins were gradually re-natured by 4 serial 1:1 dilutions of the membrane in PBS-ED and washed free of the residual guanidine HCl in fresh PBS-ED buffer. The membrane was blocked for 1 h in PBS-ED blocking buffer containing 0.05% (v/v) Nonidet-P40, 0.25% (w/v) gelatin and 3% (w/v) BSA and 100 μg of steptavidin at room temperature for 1 h in an incubator shaker. The blots were rinsed once more with PBS-ED . The biotin-labeled K-coil (1 mg/ml) was diluted to 1: 500 in PBS-ED blocking buffer (without streptavidin) and incubated with the blots at room temperature for 1 h with gentle agitation. The membrane was washed three times with PBS-ED for 10 min each wash. Streptavidin-alkaline phosphatase conjugate (Jackson ImmunoResearch Laboratory, Inc.) was diluted (1:2500; v/v) with PBS-ED blocking buffer and incubated with the blots at room temperature for 1 h with gentle agitation. The blots were washed 3 times with PBS-ED followed by a final wash with substrate buffer (100 mM Tris-HCl, pH 9.5, containing 100 mM NaCl and 5 mM $MgCl_2$). K-coil binding was visualized by the addition of nitro-blue tetrazolium chloride and 5-bromo-4-chloro-3-indolyl phosphate (NBT/BCIP) in substrate buffer. Color development was terminated by rinsing the nitrocellulose strips with de-ionized water.

CNBr Digest of PAK-pilin(met)-E-coil

CNBr cleavage of PAK-pilin(met)-E-coil peptide was carried out by incubating 0.598 μmol of the peptide in 1.0 ml of 0.1 N HCl and 59.8 μmol of CNBr for 24 h at room temperature in the dark (Gross, 1967). The reaction mixture was then diluted five fold with dH_2O and lyophilized. Final recovery of purified CNBr cleaved PAK-pilin peptide was obtained using RP-HPLC (see above).

N-terminal Sequence Analysis of E.coli-derived Peptides

An Applied Biosystems Model 470A vapor phase sequencer was used to carry out the sequential Edman degradation sequencing reactions for the first six cycles of each of the affinity purified *E.coli*- produced peptides.

C. References

Ausubel, F., Brent, R., Kingston, R., Moore, D., Seidman, J., Smith, J., and Struhl, K. (1992). Short protocols in Molecular Biology, 2nd Edition: Greene Publishing Associates & John Wiley & Sons.

Beckman Instruments, Inc. (1997a). Optima XL-I Analytical Ultracentrifuge Manual. Spinco Business Center of Beckman Instruments, Inc. Palo Alto, CA.

Beckman Instruments, Inc. (1997b). Optima XL-A/ XL-I Data Analysis Software, Version 4.0., copyright (c) 1997. Spinco Business Center of Beckman Instruments, Inc. Palo Alto, CA.

Cantor, C. R. & Schimmel, P. R. (1980). Biophysical Chemistry, Part II, p. 539, W.H. Freeman and Co., San Francisco.

Cohn, E. J. & Edsall, J. T. (1943). Proteins, Amino Acids and Peptides as Ions and Dipolar Ions, pp. 370-381, Reinhold, New York.

Craxton, M. (1991) *Methods: A Companion to Methods Enzymol.*, **3**, 20-26.

De Francesco, R., Pastore, A., Vecchio, G., and Cortese, R. (1991) *Biochemistry*, **30**, 143-147.

Ebashi, S., Wakabayashi, T., and Ebashi, F. (1971) *J. Biochem.* **69,** 441-445.

Farrow, N. A., Muhandiram, R., Singer, A.U., Pascal,S.M., Kay, C.M., Gish,G., Shoelsen, S.E., Pawson, T., Forman-Kay, J.D., and Kay, L.E.(1994) *Biochemistry* **33**, 5984-6003.

Gross, E. (1967) *Methods Enzymol.*, **11**, 238-255.

Hayes, D. B., Laue, T. & Philo, J. (1998). Sedimentation Interpretation Program, Version 1.01. Copyright (c) 1995-1998.

Heinonen, J. K. & Lahti, R. J. (1981) *Anal. Biochem.* **113,** 313-317.

Ingraham, R. H. & Hodges, R. S. (1988) *Biochemistry* **27,** 5891-5898.

Koutz, P., Davis, G. R., Stillman, C., Barringer, K., Cregg, J., and Thill, G. (1989) *Yeast*, **5**, 167-77.

Kuntz, I. D. (1971). *J. Am. Chem. Soc.* **93,** 514-516.

Laemmli, U. K. (1970) *Nature*, **277**, 680-685.

Lee, K. K., Sheth, H. B., Wong, W. Y., Sherburne, R., Paranchych, W., Hodges, R. S., Lingwood, C. A., Krivan, H., and Irvin, R. T. (1994) *Mol. Microbiol.*, **11**, 705-713.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Journal of Molecular Cloning:A Laboratory Manual, Cold Spring Harbor Laboratory,Cold spring Harbor, New York, NY.

Maxam, A. M., and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 560-564.

Monera, O.D., Sereda, T.J., Zhou, N.E., Kay, C.M. and Hodges, R.S. (1995) *J. Pept. Sci.*, **1**, 319-392.

Pace, C.N., (1986) *Methods Enzymol.* **131**, 266-279.

Pato, M. O., Mak, A. S., and Smillie, L. B. (1981) *J. Biol. Chem.* **256,** 593-601.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) "Molecular Cloning: A Laboratory Manual", 2nd Edition , Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, NY.

Schechter, Y. (1986) *J. Biol. Chem.* **261,** 66-70.

Sereda, T. J., Mant, C. T., Sonnichsen, F. D., and Hodges, R. S. (1994) *J. Chromatogr.* **676,** 139-153.

Serrano, L. and Fersht, A.R. (1989) *Nature*, 341, 296-299.

Shortel, D. (1989) *J.Biol. Chem.* **264**, 5315-5318.

Spudich, J. A. & Watts, S. (1971) *J. Biol. Chem.* **246,** 4866-4871.

Talbot, J. A. & Hodges, R. S. (1979) *J. Biol. Chem.* **254,** 3720-3723.

Towbin, H., Staehelin, T., and Gordon, J. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 4350-4354.

Tripet, B., Wagschal,K., Lavigne, P.,and R.S. Hodges. *J. Mol. Biol.* 1998, (submitted).

Tripet, B., Chana, M.and R.S. Hodges. 1999 *J. Biol Chem.*(in preparation).

Van Eyk, J. E. & Hodges, R. S. (1988) *J. Biol. Chem.* **263,** 1726-1732.

Van Eyk, J. E., Thomas, L. T., Tripet, B., Wiesner, R. J., Pearlstone, J. R., Farah, C. S., Reinach, F. C., and Hodges, R. S. (1997) *J. Biol. Chem.* **272**, 10529-10537.

Wagschal, K., Tripet, B., Lavigne, P., Mant, C.T. and R.S. Hodges. *J.Biol.Chem.* 1999 (submitted).

Weeds, A. G. & Taylor, R. S. (1975) *Nature* **257,** 54-56.

Yagi, K., Yazawa, Y., and TUsutumu, Y. (1967) *Biochem. Biophys. Res. Commun.* **29,** 331-336.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) *Gene*, **33**, 103-119.

CHAPTER III

Mapping of a Second Actin-Tropomyosin and a Second Troponin C Binding Site within the C-terminus of Troponin I, and their importance in the Ca^{2+} Dependent Regulation of Muscle Contraction.

A version of this chapter has been published: Brian Tripet, Jennifer E. Van Eyk, and Robert S. Hodges in the *J. Mol. Biol.* 1997, 271, 728-750

A. Introduction

Ca^{2+}-dependent regulation of vertebrate skeletal and cardiac muscle contraction requires the regulatory proteins troponin (Tn) and tropomyosin (Tm) to be bound to the actin thin filament (Ebashi & Endo, 1968). Troponin consists of the three protein subunits: Troponin I (TnI), Troponin T (TnT), and Troponin C (TnC). TnI, the inhibitory subunit, can bind actin and inhibit the actomyosin Mg^{2+} ATPase activity on its own (Greaser & Gergely, 1973). The inhibitory activity of TnI is potentiated in the presence of Tm, even under conditions where Tm neither binds nor inhibits the actomyosin ATPase (Perry *et al.*, 1972). The inhibitory activity of TnI can be neutralized by the formation of a 1:1 complex with TnC in the presence of calcium (Cachia *et al.*, 1986, Chong *et al.*, 1983, Perry *et al.*, 1972, Van Eyk & Hodges, 1988, Weeks & Perry, 1978). The binding of TnC to TnI apparently changes the interactions of TnI with the thin filament, such that TnI moves away from actin (Tao *et al.*, 1989, 1990). Although TnI can cause inhibition and TnC release of inhibition, full Ca^{2+}-dependent control as well as activation (increase in the ATPase activity above that of acto-S1-Tm ATPase) occurs only in the presence of the third troponin subunit, TnT (Farah *et al.*, 1994, Greaser & Gergely, 1971, Potter *et al.*, 1995). TnT binds tropomyosin as well as TnI and TnC and is believed to be the "transmission link" between the conformational changes occurring between the TnI•TnC complex and Tm.

In order to understand at a molecular level how the binding of Ca^{2+} to TnC leads to the eventual movement of Tm proposed in the steric blocking model, extensive structure/function studies using a variety of techniques have been carried out to determine the structure, function and sites of interactions between each of the troponin components and their change in conformation upon binding calcium (for recent reviews see (Farah & Reinach, 1995, Grabarek *et al.*, 1992, Zot & Potter, 1987)). Structural investigations into TnC, the Ca^{2+} binding component of troponin, has revealed that it exists as a dumbbell shaped molecule with two globular domains linked by a long central helix (Herzberg & James, 1985, Sundaralingam *et al.*, 1985). Each domain possesses two metal ion binding sites of the EF hand structural motif (Collins *et al.*, 1973, Kretsinger & Nockolds, 1973). The two metal binding sites located within the NH_2-terminal domain, sites I and II, are referred to as the "regulatory sites", as they are lower in Ca^{2+} affinity than those located within the C-domain, but are critical for the Ca^{2+}-dependent regulation of muscle contraction ((Farah *et al.*, 1994, Grabarek *et al.*, 1992, Potter *et al.*, 1995) and references therein). The two metal binding sites, sites III and IV, located within the C-domain are referred to as the "high affinity" sites which can bind both Ca^{2+} and Mg^{2+} ions competitively ((Farah *et al.*, 1994, Grabarek *et al.*, 1992, Potter *et al.*, 1995) and references therein). In the crystal structure, sites III and IV are bound with calcium while those within the N-domain are free. Thus, it is believed that the crystal structure may represent the structure of TnC in the "relaxed state" of muscle regulation.

In an attempt to understand the initial conformational change created by the binding of Ca^{2+} to the N-domain of TnC and thus the Ca^{2+}-induced trigger of muscle contraction, Hertzberg, Moult and James proposed a model for the conformation of the N-domain of TnC based on its Ca^{2+} saturated C-domain (Herzberg *et al.*, 1986). In this model, the authors proposed that upon binding Ca^{2+} ions, the regulatory sites I and II would alter the spatial relationship of the helices within the N-domain (N, B and C helices relative to helices A and D) to open up and expose partially and totally buried nonpolar residues and thus form a "hydrophobic pocket". This newly exposed site is believed to be a new binding site for TnI, although the residues of TnI which bind this site have not been definitively defined. Recent studies have now given strong experimental support to this model (Fujimori *et al.*, 1990, Gagne *et al.*, 1994, 1995, Grabarek *et al.*, 1990, Sheng *et al.*, 1990, Slupsky *et al.*, 1995, Slupsky & Sykes, 1995).

At present, no high resolution structure for skeletal TnI is yet available, although a significant amount of topological and functional information is known about its interactions with TnC, TnT and actin-Tm. Syska *et al.* (Syska *et al.*, 1976) were the first to identify that the inhibitory activity of TnI is carried out by only a small section of the complete protein, residues 96-116, often referred to as the "inhibitory region". Studies into this region by Talbot and Hodges using synthetic peptides revealed that residues 104-115 represent the minimal sequence required to cause inhibition (Talbot & Hodges, 1981b). Subsequent studies have since shown that a synthetic peptide, $TnI_{104-115}$, can bind to actin-Tm and inhibit the acto-S1-Tm ATPase activity as well as be neutralized by TnC in a Ca^{2+}-dependent manner (Van Eyk *et al.*, 1993). Furthermore, Van Eyk *et al.* (1993) have shown that this same peptide ($TnI_{104-115}$) can replace intact TnI in regulating the Ca^{2+}-dependent contraction and relaxation of skinned muscle fibers, thus suggesting that the major switch between contraction and relaxation involves a switch of the inhibitory region of TnI between actin-Tm and TnC.

In addition to the inhibitory region, other regions of TnI have also been identified as important in the regulatory process. TnI residues ~40-96 located towards the N-terminus of the inhibitory region have been shown to interact with TnT and are now believed to be important for the "activation" of the ATPase activity which is observed only in the presence of TnT (Chong & Hodges, 1982, Farah *et al.*, 1994, Hichcock, 1981, Potter *et al.*, 1995, Sheng *et al.*, 1992). TnI residues ~1-40 (the N-terminus) have been shown to bind the C-domain of TnC and are believed to play an important structural role in anchoring TnC to the ternary troponin complex in the absence of calcium. For example, a TnI mutant in which the N-terminal residues 1-57 have been deleted (TnI_{del57}) shows a modified interaction with TnT and is unable to form a ternary complex with TnT and TnC in the absence of calcium (Sheng *et al.*, 1992). Functional studies of this region by Ngai and Hodges (1992) using synthetic peptides have shown that a TnI peptide comprising residues 1-40 of the N-terminus can block the neutralization of the inhibitory activity of TnI or $\text{TnI}_{104\text{-}115}$ by TnC, suggesting that the N-terminus may play both a structural and functional role.

Although significantly less attention has been given to the C-terminal residues (116 to 182) of TnI, recent studies suggest these residues play significant roles in the inhibitory activity of TnI and its interactions with TnC. For example, investigations with recombinantly expressed TnI fragments have shown that fragments which lack either the first 57 (TnI_{d57}), 102 ($\text{TnI}_{103\text{-}182}$) or COOH-terminal 26 residues ($\text{TnI}_{1\text{-}156}$) are able to retain similar inhibitory activities to that of the intact TnI protein (Farah *et al.*, 1994, Sheng *et al.*, 1992). Fragments which lack the C-terminal residues ($\text{TnI}_{1\text{-}116}$) or just encompass the inhibitory region ($\text{TnI}_{96\text{-}115}$ or $\text{TnI}_{104\text{-}115}$) have significantly less inhibitory activity (on a molar basis) when compared to intact TnI. Farah *et al.* (1994) have suggested that it is residues within the region of 116 to 156 which are responsible for this effect. More recently, however, Van Eyk *et al.* (1997) have reported that a TnI fragment encompassing only residues 96-148 ($\text{TnI}_{96\text{-}148}$) is also able to elicit an equivalent inhibitory activity to that of the intact TnI protein, thus further delineating the minimal region required for "wild type like" inhibition to residues 96-148 of TnI and the residues responsible for this effect to the region 116 to 148.

In addition to the enhanced inhibitory activity, recent experimental studies have also indicated that residues located within the C-terminus of TnI may also be involved in interacting with the regulatory domain (N-domain) of TnC. For example, Kobayashi *et al.* (1991) have demonstrated that TnC mutants containing single cysteine residues at position 12 or 57 of rabbit skeletal TnC and modified with 4-maleimidobenzophenone can be crosslinked to TnI in the region 113-121 and 132-141, respectively. Furthermore, fluorescence studies by Pearlstone and Smillie (1995a) using recombinant TnC mutants with single tryptophan residues within the N and C-domains of TnC, have indicated that a recombinant TnI fragment consisting of residues 96-148 contains a significantly greater affinity for TnC than the individual TnI fragments of 96-116 or 117-148 alone, as well as observing that the C-terminus of this region can significantly enhance the affinity for just the N-domain of TnC (Pearlstone and Smillie 1995a, 1997).

Thus, in the present study, we have decided to focus on mapping and identifying the specific residues or regions of residues within the C-terminus of TnI (residues 116 to 148) that also appear to be involved in inhibition and TnC binding. In this respect, we have prepared several synthetic TnI peptide analogs containing truncations, substitutions and modifications within the C-terminal region 96-148 of rabbit skeletal TnI, and subsequently analyzed these analogs in reconstituted thin filament inhibitory, binding, and release assays in order to determine their effects.

B. Results

Identifying the Critical Residues within TnI_{96-148} involved in Inhibition

To identify the critical residues located within the C-terminal region (residues 116-148) of TnI_{96-148} that appear to be also involved in the inhibitory activity of TnI, several synthetic peptide analogs of the region 96-148 of rabbit skeletal TnI were prepared (see Table III-1 for sequences) and analyzed for their ability to inhibit the acto-S1-Tm ATPase activity. (Fig. III-1, panels A, B, and C (inset)).

Synthetic TnI peptide TnI_{96-148} inhibits the ATPase activity in a concentration-dependent manner that is indistinguishable from that of intact TnI protein. This peptide inhibits 85% of the ATPase activity at a 1:1 mole ratio of inhibitor:Tm. The inhibitory activity of this peptide is significantly greater than that of TnI peptides, $TnI_{104-115}$ and TnI_{96-115}, which encompass only the primary inhibitory region. These peptides require a 6:1 and 2:1 mole ratio of inhibitor:Tm, respectively, to achieve the same inhibitory maximum (compare the inhibition curves for TnI_{96-115}, $TnI_{104-115}$ and TnI_{96-148}, Fig. III-1, panel A). Truncation of the C-terminal 9 residues (140-148) from the TnI region 96-148 causes a dramatic decrease in the ability of the remaining peptide (TnI$_{96-139}$) to inhibit the acto-S1-Tm ATPase activity (Fig. III-1A). This significant decrease suggests that residues important for actin-Tm binding and/or inhibition have now been deleted. Truncation of the next 8 C-terminal residues (TnI$_{96-131}$) causes no further effect (Fig III-1., panel A; compare TnI$_{96-139}$ and TnI$_{96-131}$) Surprisingly, when the next 16 residues are removed (residues 116-131), leaving only the inhibitory region, the inhibitory activity of this peptide (TnI$_{96-115}$) now appears to increase relative to that of the latter two peptides (compare TnI$_{96-115}$ to TnI$_{96-131}$ and TnI$_{96-139}$). The decreased inhibitory activity observed for TnI$_{96-139}$ and TnI$_{96-131}$ relative to that of TnI$_{96-115}$ would suggest that residues 116-139 are not important for the actin-Tm inhibitory interaction and that their presence in the absence of the C-terminal 9 residues (140-148) only entropically hampers binding of the inhibitory region (residues 96-115).

To investigate further the importance of the central residues (116-131) and C-terminal residues (140-148), we analyzed the inhibitory activity of the TnI analogs which contained substitutions or modifications within these regions (TnI$_{96-148A4-A8}$, for sequences see Table III-1; Fig. III-1, panel 1B). Substitution of the two leucine residues (residues 125 and 126) (TnI$_{96-148A5}$) within the central region (116-131) to alanines causes no effect upon the inhibitory activity. Similarly, when the three methionine residues within the central region (residues 116, 121 and 134) were oxidized (methionine sulfoxide), this peptide (TnI$_{96-148A1}$) showed an identical inhibition curve to that of TnI$_{96-148}$ in the reduced state, suggesting that the methionine residues are not involved in inhibition. Substitution of methionine 121 and leucine 122 (TnI$_{96-148A6}$), did cause a decrease in the inhibitory activity, such that the peptide now exhibited inhibitory activity equivalent to that of TnI$_{96-115}$. Since oxidation of methionine 121 did not appear to cause an effect, this would suggest that leucine 122 is important for inhibition. Surprisingly, this same small decrease in the inhibitory activity was observed when all 10 of the residues between 116-126 were replaced by glycine residues (compare TnI$_{96-148A7}$ with TnI$_{96-148A1}$).

Table III-1. Amino acid sequences of the synthetic troponin I peptides

Peptide[a]	Abbreviated name[b]	Sequence[c]
		96 115 131 139 148
TnI 96-148(C133A,T148G)	TnI$_{96-148}$	Ac-NQKLFDLRGKFKRPPLRRVRMSADAMLKALLGSKHKVAMDLRANLKQVKKEDG-amide
TnI 96-139(C133A)	TnI$_{96-139}$	Ac-NQKLFDLRGKFKRPPLRRVRMSADAMLKALLGSKHKVAMDLRAN-amide
TnI 96-131	TnI$_{96-131}$	Ac-NQKLFDLRGKFKRPPLRRVRMSADAMLKALLGSKHK-amide
TnI 96-115	TnI$_{96-115}$	Ac-NQKLFDLRGKFKRPPLRRVR-amide
TnI 96-115(F106G,R115G)	TnI$_{96-115A}$	Ac-NQKLFDLRGK**G**KRPPLRRV**G**-amide
TnI 104-115	TnI$_{104-115}$	Ac-GKFKRPPLRRVR-amide
TnI 115-131	TnI$_{115-131}$	Ac-RMSADAMLKALLGSKHK-amide
TnI 128-148(C133A,T148G)	TnI$_{128-148}$	Ac-SKHKVAMDLRANLKQVKKEDG-amide
TnI 128-148(C133A,K(141,144,145)A,T148G)	TnI$_{128-148A}$	Ac-SKHKVAMDLRANL**A**QV**AA**EDG-amide
TnI 96-148(M(116,121,134)M(O),C133A,T148G)	TnI$_{96-148A1}$	Ac-NQKLFDLRGKFKRPPLRRVR$\overset{*}{\mathbf{M}}$SADA$\overset{*}{\mathbf{M}}$LKALLGSKHKVA$\overset{*}{\mathbf{M}}$DLRANLKQVKKEDG-amide
TnI 96-148(C133A,K(141,144,145)A,T148G)	TnI$_{96-148A2}$	Ac-NQKLFDLRGKFKRPPLRRVRMSADAMLKALLGSKHKVAMDLRANL**A**QV**AA**EDG-amide
TnI 96-148(F106G,R115G,C133A,T148G)	TnI$_{96-148A3}$	Ac-NQKLFDLRGK**G**KRPPLRRV**G**MSADAMLKALLGSKHKVAMDLRANLKQVKKEDG-amide
TnI 96-148(F106G,R115G,C133A, K(141,144,145)A,T148G)	TnI$_{96-148A4}$	Ac-NQKLFDLRGK**G**KRPPLRRV**G**MSADAMLKALLGSKHKVAMDLRANL**A**QV**AA**EDG-amide
TnI 96-148(L125A,126A,C133A,T148G)	TnI$_{96-148A5}$	Ac-NQKLFDLRGKFKRPPLRRVRMSADAMLKA**AA**GSKHKVAMDLRANLKQVKKEDG-amide
TnI 96-148(M121A,L122A,C133A,T148G)	TnI$_{96-148A6}$	Ac-NQKLFDLRGKFKRPPLRRVRMSADA**AA**KALLGSKHKVAMDLRANLKQVKKEDG-amide
TnI 96-148(G116-126,C133A,T148G)	TnI$_{96-148A7}$	Ac-NQKLFDLRGKFKRPPLRRVR**GGGGGGGGGGG**GSKHKVAMDLRANLKQVKKEDG-amide
TnI 96-148(G116-126,C133A,M134M(O),T148G)	TnI$_{96-148A8}$	Ac-NQKLFDLRGKFKRPPLRRVR**GGGGGGGGGGG**GSKHKVA$\overset{*}{\mathbf{M}}$DLRANLKQVKKEDG-amide
TnI 1-40	Rp$_{1-40}$	Ac-GDEEKRNRAITARRQHLKSVMLQIAATELEKEEGRREAEK-amide
Control peptide[d]	control peptide	Ac-IQEAKEILVEHYDNIEQKIDDIDHEIADLQAKRTRLVQQHPRIDE-amide

---

[a] Peptide sequence name followed by residues (indicated in brackets) substituted from that of native rabbit skeletal TnI. For example, an analog with threonine 148 of native TnI substituted by glycine is denoted T148G. M(O) denotes methionine sulfoxide.

[b] Abbreviated name of the peptide used within the text. The subscript "A" denotes analog and the number specifies a specific analog as indicated.

[c] Amino acid sequence of the peptides. Methionines that exist as methionine sulfoxide are indicated with a star (*). The key residues substituted are denoted in bold and underlined. Ac- denotes N-terminal acetylation, -amide denotes C-terminal amide.

[d] The control peptide sequence is that of the C-terminus of the E.*coli* ProP protein (residues 450-500), an unrelated protein.

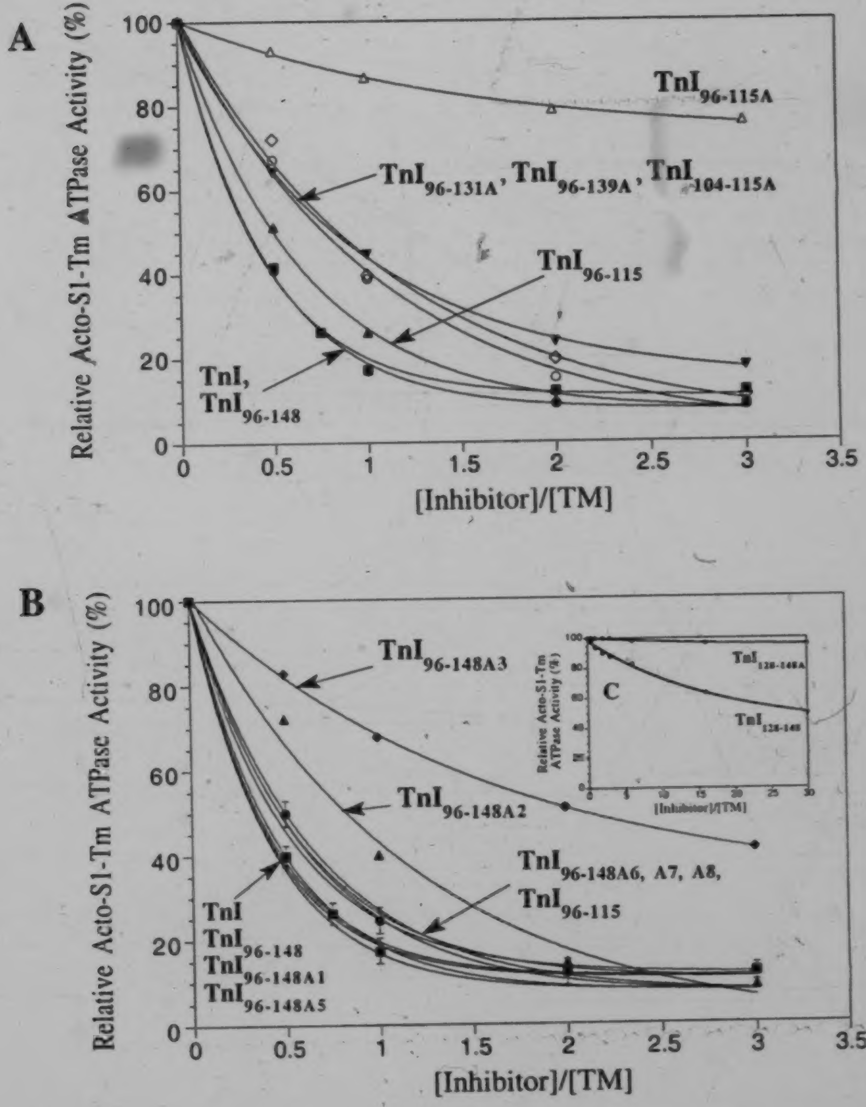


Figure III-1. Inhibition of the acto-S1-Tm ATPase activity by TnI and TnI peptides. The acto-S1-Tm ATPase activity was measured as a function of increasing quantities of TnI or TnI peptides as described in "Materials and Methods". Panel A: native rabbit skeletal TnI (■); TnI$_{96-148}$ (●); TnI$_{96-139}$ (O); TnI$_{96-131}$ (◊); TnI$_{96-115}$ (▲); TnI$_{96-115A}$ (Δ); TnI$_{104-115}$ (▼). Panel B: TnI, TnI$_{96-148}$,TnI$_{96-148A1}$, TnI$_{96-148A5}$ (■); TnI$_{96-115}$,TnI$_{96-148A6}$, TnI$_{96-148A7}$, TnI$_{96-148A8}$ (●); TnI$_{96-148A2}$ (▲); TnI$_{96-148A3}$ (◆). Panel C (inset): TnI$_{128-148}$ (●); TnI$_{128-148A}$ (O); In Panel B,similar inhibition curves are indicated by a single symbol with indicated arrow bars.

3

The fact that no further decrease in the inhibitory activity occurs upon replacement of the 10 central residues suggests that these residues 116-126 (with the exception of leucine 122) are not involved in the interactions required for inhibition. In contrast to the central region of TnI$_{96-148}$, when substitutions were made within the C-terminus (140-148), significant losses in the inhibitory activity were observed. Substitution of the three terminal lysine residues (K$_{141}$, K$_{144}$ and K$_{145}$) located within the C-terminal region 140-148 causes the TnI peptide (TnI$_{96-148A2}$) to show a similar inhibitory activity to that observed for the truncated TnI peptides TnI$_{96-139}$ and TnI$_{96-131}$. The significant decrease in inhibitory activity observed only upon substitution or deletion of the C-terminus of the TnI 96-148 region suggests that a second inhibitory interaction site exists within the C-terminus of TnI (140-148), and is non-contiguous with the inhibitory region (96-115).

To determine whether the C-terminal region containing residues 140-148 possesses any inhibitory activity, we prepared two synthetic TnI peptides of the C-terminus from residues 128-148 and determined their effect on the acto-S1-Tm activity. Addition of TnI$_{128-148}$ (Fig. III-1, panel C (inset)) induces a weak but significant inhibitory effect on the acto-S1-Tm activity (~ 10% at a 3:1 ratio of inhibitor to Tm) in the same concentration range as that measured for the other TnI analogs. If increasing amounts of the peptide are added, inhibition can be increased to ~ 50% at a 30:1 peptide:Tm mole ratio. Substitution of the three terminal lysine residues (K$_{141}$, K$_{144}$, and K$_{145}$) with alanine residues completely abolishes this effect (TnI$_{128-148A}$), similar to that seen above for the TnI$_{96-148A2}$ peptide with the same substitutions. The dramatic effect of the lysine to alanine residue substitutions illustrates again the importance of the three terminal lysine residues in the second actin-Tm interaction site.

Since it appeared as if the C-terminal region (residues 128-148) could induce a weak inhibitory response on its own, we next investigated whether a combination of the two TnI regions 96-115 and 128-148 would have a synergistic effect on the inhibitory activity and generate the same inhibitory response to that of TnI$_{96-148}$. Results of these experiments (data not shown) only showed degrees of inhibition similar to that of TnI_{96-115} alone (at equivalent concentrations) irrespective of the presence of the $TnI_{128-148}$ peptide (this peptide was added up to a mole ratio of 1:15 of TnI_{96-115} to $TnI_{128-148}$). Thus, the two regions apparently must be connected to each other in order to achieve maximum inhibitory activity.

Lastly, although the effects of substituting or deleting the three terminal lysine residues (K141, L144 and K145) are significant, they are not as significant as when residues are substituted or deleted from the primary inhibitory region (residues 96-115) itself. Truncation of the 8 N-terminal residues (96-103) to generate the smallest inhibitory peptide ($TnI_{104-115}$), causes a dramatic decrease in the inhibitory activity. In fact, $TnI_{104-115}$ now inhibits to a degree equivalent to that of the two C-terminally truncated peptides TnI_{96-131} and TnI_{96-139}. This significant decrease indicates that the 8 N-terminal residues are playing an important role in the inhibition between this region of TnI and the thin filament. Substitution of residues phenylalanine 106 and arginine 115 by glycine almost completely abolishes the inhibitory activity (compare TnI_{96-115} and $TnI_{96-115A}$, Fig. III-1A). Interestingly, when the same substitutions are made within TnI_{96-148}, the inhibitory activity is also decreased, but not to the same extent (compare the inhibitory activities between $TnI_{95-115A}$ (panel A) and $TnI_{96-148A3}$ (panel B)). The greater inhibitory activity observed in the presence of the C-terminal residues (116-148) supports the above observation that the C-terminal residues contain a second inhibitory site (140-148).

Actin-Tm Centrifugation Studies

To establish whether differences in the inhibitory activities (Fig. III-1) can be correlated with changes in the ability of the various TnI peptides to bind the actin-Tm filament, we tested the ability of each of the TnI peptides to co-sediment with actin-Tm in a quantitative sedimentation assay. From Fig. III-2, the inhibitory region (residues 96-115) appears to be the dominant site of interaction between this region of TnI and actin-Tm. All of the TnI

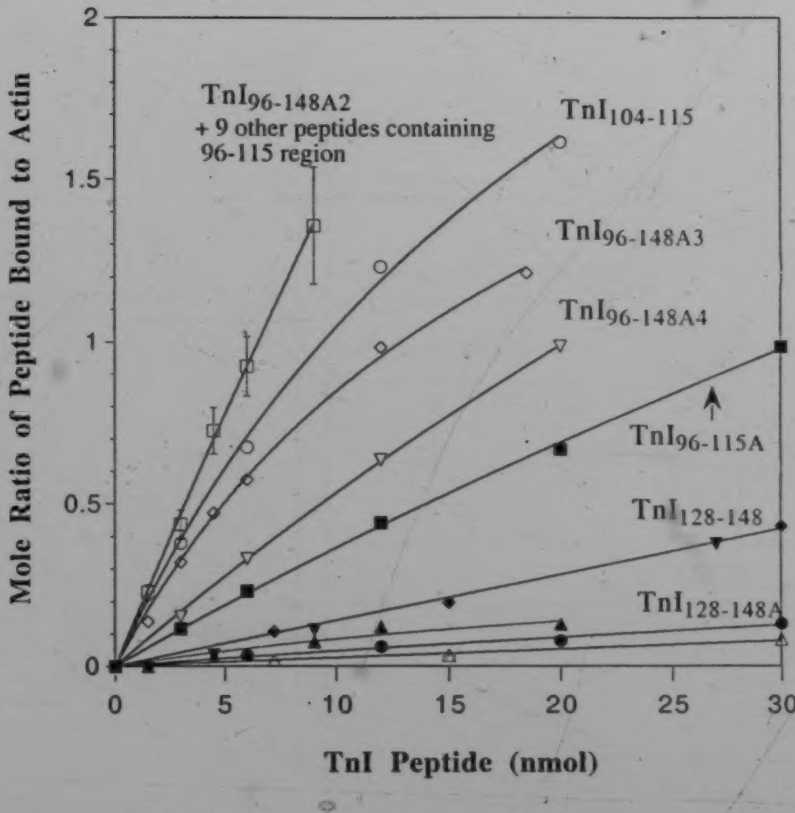


Figure III-2. Actin-Tm centrifugation studies. The quantity of TnI peptides bound to the thin filament at various peptide concentrations was measured, and the data plotted as a mole ratio of peptide bound to actin. TnI$_{96-115}$,TnI$_{96-131}$,TnI$_{96-139}$,TnI$_{96-148}$, TnI$_{96-148A1}$, TnI$_{96-148A2}$,TnI$_{96-148A5}$,TnI$_{96-148A6}$,TnI$_{96-148A7}$, TnI$_{96-148A8}$ are indicated by the single symbol (❑) with associated error bars due to their binding equivalence; TnI$_{104-115}$ (O); TnI$_{96-148A3}$ (◊); TnI$_{96-148A4}$ (▽); TnI$_{96-115A}$ (■); TnI$_{128-148}$ (◆); TnI$_{128-148}$ in the presence of 3 nmoles of TnI$_{96-115}$ (▼); Rp$_{1-40}$ (▲); TnI$_{128-148A}$ (●); Control Peptide (Δ).

peptides which contain this region appear to bind similarly to the actin-Tm filament . Although the inhibition assays showed differences amongst these peptides, particularly in regards to the TnI peptides where the C-terminal 9 residues (140-148) were either deleted or substituted (Fig. III-1), either the sensitivity of this assay is not able to discriminate amongst these small differences or that these peptides have similar binding affinities for the thin filament and the loss of inhibitory activity is due to a loss of efficacy.

Only when residues are deleted or substituted within the inhibitory region are significant differences in the binding affinities observed. For example, truncation of the 8 N-terminal residues (96-103) to generate the smallest inhibitory peptide (TnI$_{104\text{-}115}$) results in a weaker binding affinity than that of TnI$_{96\text{-}115}$; thus the decreased inhibitory activity of TnI$_{104\text{-}115}$ compared to TnI$_{96\text{-}115}$ (Fig. III-1) can be ascribed to a decreased binding affinity of TnI$_{104\text{-}115}$ for actin-Tm filament. Similarly, substitution of phenylalanine 106 and arginine 115 with glycine residues within the inhibitory region also now causes a dramatic decrease in the ability of these peptides to bind to the actin-Tm filament (TnI$_{96\text{-}115A}$, TnI$_{96\text{-}148A3}$, Fig. III-2), which also correlates well with the observed substantial decrease in the ability of these peptides to inhibit the acto-S1-Tm activity (Fig. III-1, panel A and B). Interestingly, only when the binding affinity of the inhibitory region is reduced by the phenylalanine 106 and arginine 115 substitutions can the binding interaction between the C-terminus and the thin filament be seen more clearly, (compare TnI$_{96\text{-}148A2}$ to TnI$_{96\text{-}148A3}$). When the three C-terminal lysine residues (141, 144, and 145) are substituted with alanines, a further reduction in the ability of TnI$_{96\text{-}148A4}$ peptide to bind actin-Tm is observed (compare TnI$_{96\text{-}148A3}$ to TnI$_{96\text{-}148A4}$). Thus, the lack of inhibitory activity observed for peptide TnI$_{96\text{-}148A4}$ compared to TnI$_{96\text{-}148A3}$ (Fig. III-1, panel B) can also be ascribed to an additional loss in actin-Tm binding. Similarly, TnI peptide 128-148 which showed only a weak inhibitory activity, also binds only weakly to actin-Tm, and this binding is also completely abolished by the substitutions of the C-terminal lysines (compare TnI$_{128\text{-}148}$ to TnI$_{128\text{-}148A}$ and TnI$_{96\text{-}148A3}$ to TnI$_{96\text{-}148A4}$). The fact that TnI$_{128\text{-}148}$ binds, albeit weakly, under these conditions further supports the idea that a second actin-Tm binding site does exists within this region of TnI.

When $TnI_{128-148}$ and TnI_{96-115} peptides were both present in the assay (as described above for the evaluation of synergistic inhibition), $TnI_{128-148}$ showed an identical binding curve to that observed in the absence of TnI_{96-115}, suggesting that the two binding regions within the TnI 96-148 region are binding two separate sites on the actin-Tm filament.

Because of the observation of greater than 1:1 binding and its implications, we re-examined whether the observed binding of the inhibitory region peptides to actin was strictly due to specific binding and not a result of non-specific binding due to the highly basic nature of the inhibitory region. In this regard, we analyzed two control peptides, one of which is the highly basic N-terminal region of TnI (residues 1-40) which is known not to bind actin (Ngai & Hodges, 1992, Syska *et al.*, 1976) and an un-related control peptide. Both of these peptides did not co-sediment with the thin filament at all (Fig. III-2, peptides Rp 1-40 and control) under any conditions. Furthermore, under these conditions, all of the TnI peptides which did co-sediment with actin-Tm did so only in the presence of actin-Tm (i.e., peptides alone do not sediment (data not shown)). Thus, the binding interactions observed appear to be a result of specific interactions with the filament.

Neutralization of the Inhibitory Activity of TnI and TnI Peptides by TnC (plus Ca^{2+})

In addition to the enhanced inhibitory activity observed for the TnI region 96-148, we addressed the question of a second possible site for TnC interaction in this region. To determine the approximate location of this site and the amino acid residues involved, the ability of TnC to neutralize the inhibition induced by the various TnI peptides were investigated. A 2:1 mole ratio inhibitor:Tm was used, conditions that produce maximal inhibition in the absence of TnC for all peptides. Increasing quantities of TnC (plus Ca^{2+}) were added up to a TnC :inhibitor ratio of 4:1 (Fig. III-3A). Neutralization of the synthetic

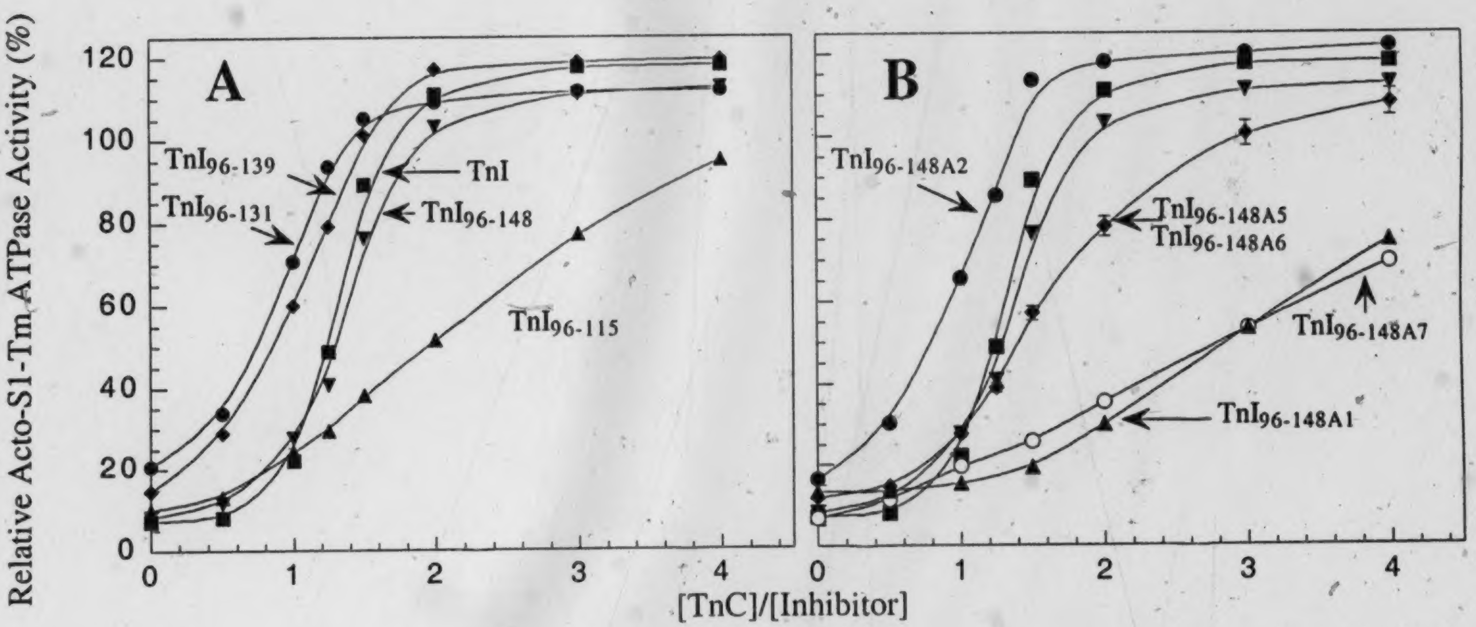


Figure III-3. Neutralization of TnI or TnI peptide inhibition by TnC (plus Ca^{2+}). The ability of TnC (in the presence of 0.4 mM $CaCl_2$) to neutralize the inhibitory effect of TnI or TnI peptides was measured by the increase in ATPase activity with increasing concentrations of TnC. All inhibitors were present at a 2:1 mole ratio of inhibitor to Tm. Panel A: TnI (■); TnI_{96-148} (▼); TnI_{96-139} (◆); TnI_{96-131} (●); TnI_{96-115} (▲). Panel B: TnI (■); TnI_{96-148} (▼); $TnI_{96-148A2}$ (●); $TnI_{96-148A5}$ and $TnI_{96-148A6}$ (◆); $TnI_{96-148A1}$ (▲); $TnI_{96-148A7}$ (O).

TnI_{96-148} peptide by TnC shows a similar release curve to that observed for intact TnI protein (Fig. III-3A). Truncation of the C-terminal nine residues (TnI_{96-139}) results in a shift in the release curve to lower concentrations of TnC required for full release. This shift, however, is in agreement with the idea that the C-terminal nine amino acid residues contain (or partly encompass) a second actin-Tm binding site. Thus deletion of these residues reduces the peptide's effective actin-Tm binding affinity/interaction and a lower concentration of TnC is required to release inhibition of the acto-S1-Tm ATPase activity. Truncation of the next 8 amino acid residues (TnI_{96-131}) also shows a similar shift in the release curve. However, removal of the next 16 residues (TnI_{96-115}) causes a dramatic loss in the ability of TnC to neutralize the inhibitory activity of this peptide. This significant decrease would suggest that a site important for TnC interaction has been deleted. Although TnI_{96-115} is a better inhibitor than TnI_{96-131} or TnI_{96-139}, it is weaker than that of TnI_{96-148} (Fig. III-1, panels A and B) which is still able to be fully released, thus dismissing the idea that the observed differences in release are a result of their inhibitory activities.

To establish that the second actin-Tm binding site is in fact responsible for the shifts observed for TnI_{96-139} and TnI_{96-131} relative to that of TnI_{96-148}, we analyzed the TnI peptide analog $TnI_{96-148A2}$ which contains the three C-terminal lysine to alanine substitutions (Fig. III-3B). This peptide now also showed a TnC release curve that was shifted to a lower TnC concentration, similar to those of TnI_{96-131} and TnI_{96-139}, indicating that deletion or substitution of residues within the "second actin-Tm binding site" are responsible for the shift in the release assay.

To investigate further the importance of residues 116-131 of TnI in the ability of TnC to neutralize the inhibitory activity, we analyzed the TnI peptide analogs in which we had substituted or modified residues within this region. Fig. III-3, panel B shows that substitution of the hydrophobic residues (methionine 121, leucine 122 ($TnI_{96-148A6}$) and leucine 125, leucine 126 ($TnI_{96-148A5}$)) with alanines results in a decreased ability of TnC to release the inhibitory activity for these peptides compared to that of the unsubstituted analogs, although full release can still occur but at a higher TnC:peptide ratio (compare $TnI_{96-148A5}$ and $TnI_{96-148A6}$ to TnI_{96-148}). It is interesting that both substitution analogs show a similar effect (i.e., the release curves appear to be superimposable), suggesting that both sets of hydrophobes are playing equivalent roles in the interaction with TnC. When all 10 of the residues between 116-126 are replaced with glycine residues, the TnI peptide ($TnI_{96-148A7}$) now drastically loses its ability to be released, in a manner similar to that seen for the deletion of these same residues (compare $TnI_{96-148A7}$ with TnI_{96-115}). This result would strongly indicate that it is the residues between 116-126 that are critical for this interaction with TnC. Surprisingly, the $TnI_{96-148A1}$ analog which contains the three oxidized methionine residues also shows the same significant effect on the ability of TnC to release the inhibitory activity in a manner similar to that of the 10-residue glycine substitution. This would suggest that either methionine 116 or 121 or both are very important for their interaction with TnC. Although there is a third methionine in the C-terminal region (residue 134), the TnI analog $TnI_{96-148A8}$ peptide which contains this residue oxidized did not show any further decrease in its ability to be released by TnC compared to that of its reduced state $TnI_{96-148A7}$ (data not shown).

Since both oxidation of methionine residues and substitution with the glycine residues is predominantly affecting only the hydrophobic residues within the region 116-131, this would suggest that the important interactions between TnC and this region of TnI are of a hydrophobic nature. Furthermore, the fact that the release curves for peptides $TnI_{96-148A1}$ and $TnI_{96-148A7}$ are similar to that of TnI_{96-115} indicates that the interaction between the inhibitory region and TnC has not changed. As previously reported, it is believed that the inhibitory region interacts with the C-domain of TnC (Cachia *et al.*, 1983, 1986, Campbell *et al.*, 1991, Campbell & Sykes, 1991, Chandra *et al.*, 1994, Dalgarno *et al.*, 1982, Drabikowski *et al.*, 1985, Grabarek *et al.*, 1981, Howarth *et al.*, 1995, Jha *et al.*, 1996, Kobayashi *et al.*, 1994, 1996, Krudy *et al.*, 1994, Lan *et al.*, 1989, Leszyk *et al.*, 1987, 1988, Ngai *et al.*, 1994, Slupsky *et al.*, 1992, Swenson & Fredricksen, 1992, Tsuda *et* *al.*, 1992, Wang *et al.*, 1990, Weeks & Perry, 1978) and that the N-domain, upon binding calcium, exposes a new hydrophobic surface area (Gagne *et al.*, 1994,1995, Herzberg *et al.*, 1986, Slupsky *et al.*, 1995, Slupsky & Sykes, 1995). With the above results indicating that TnI residues 116-131 interact with TnC in a hydrophobic manner, these results strongly suggest that it is these residues that are interacting with the N-domain of TnC.

One further point that can be made regarding the TnC release results. We believe the sigmoidal shape, normally suggestive of cooperativity, is not a result of cooperativity but rather a result of the assay conditions themselves. In both the TnC and T·C release assays (see below), a 2:1 mole ratio of inhibitor:Tm was present in the assay, which for some peptides exceeds the amount required for maximal inhibition (see Fig. III-1). The initial lag is then a result of the initial titration of this excess inhibitor. Inhibitor conditions could have been chosen for each peptide or protein individually at their maximal point of inhibition; however, for a direct comparison of all of the TnI peptides with TnI protein, a 2:1 mole ratio was used throughout. Thus, in this assay, we are directly comparing the relative differences in neutralization of the various TnI inhibitors at the same concentrations.

Effect of T·C Complex on the Inhibitory Activity of TnI and TnI Peptides (plus Ca^{2+})

To verify that the observed effects between the TnI peptides and TnC (Fig. III-3) also occur in the presence of TnT, the third subunit of the troponin complex, we re-investigated the neutralization of the inhibitory activity of the various TnI peptides by T·C complex (plus Ca^{2+}) (Fig. III-4). Overall the results with T•C complex are similar to that observed for TnC neutralization of the TnI peptides, with the exception that intact TnI can activate the ATPase activity above that of acto-S1-Tm levels (Fig. III-4, panel C, inset). Fig. III-4, panel A, shows that truncation of the C-terminal 9 and 17 residues from the TnI 96-148 region (TnI peptides TnI_{96-139} and TnI_{96-131}, respectively) results in full release of the acto-S1-Tm inhibition; however, the release curves are shifted to lower T·C complex concentrations

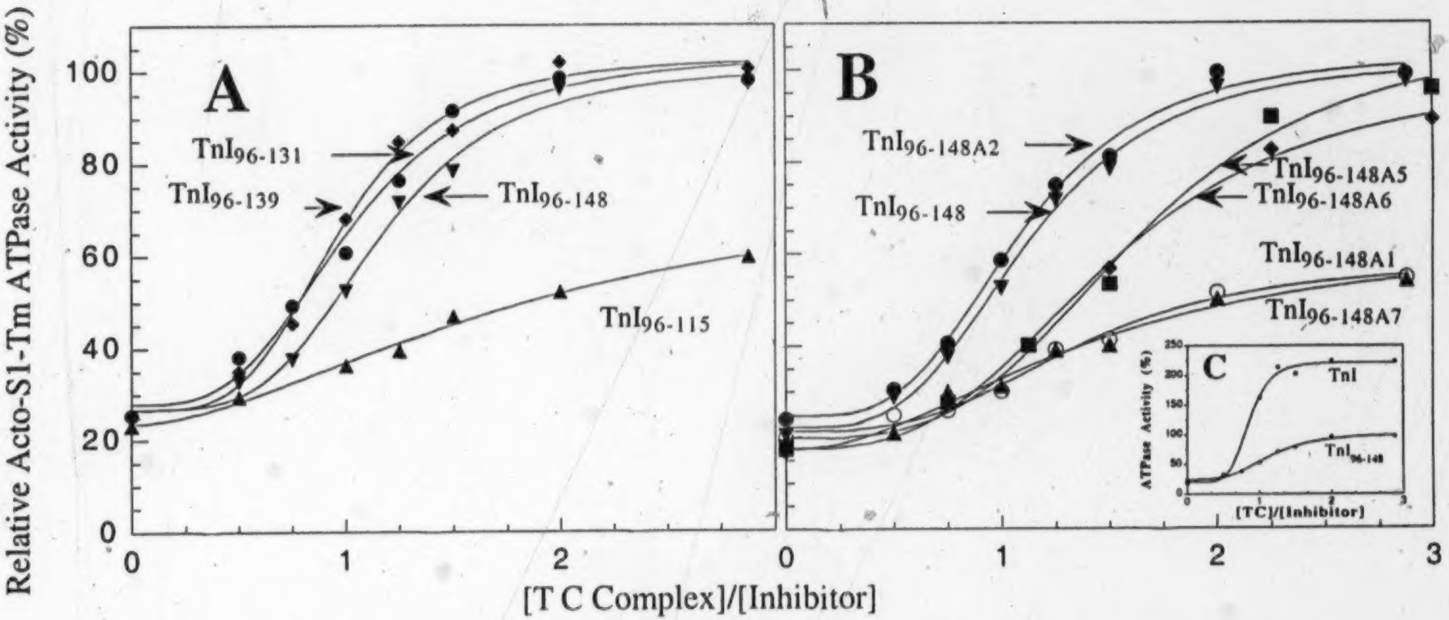


Figure III-4. Neutralization of TnI or TnI peptide inhibition by T·C complex (plus Ca^{2+}). The ability of T·C complex (in the presence of 0.4 mM $CaCl_2$) to neutralize the inhibitory effect of TnI or TnI peptides was measured by the increase in ATPase activity with increasing concentrations of T·C complex. All inhibitors were present at a 2:1 mole ratio of inhibitor to Tm. Panel A: TnI_{96-148} (▼); TnI_{96-139} (●); TnI_{96-131} (◆); TnI_{96-115} (▲). Panel B: TnI_{96-148} (▼); $TnI_{96-148A2}$ (●); $TnI_{96-148A5}$ (■);$TnI_{96-148A6}$ (◆); $TnI_{96-148A1}$ (O); $TnI_{96-148A7}$ (▲). Panel C: TnI (O);TnI_{96-148} (▼).

than that required for TnI_{96-148}. The sizes of these apparent shifts are smaller than those previously observed (Fig. III-3) for these same peptides, and may be due to modulation caused by TnT interaction. Substitution of the three C-terminal lysine residues ($TnI_{96-148A2}$), which disrupts the second actin-Tm binding site, also shifts the T·C release curve slightly to that of lower T•C complex concentrations relative to that observed for TnI_{96-148} (Fig. III-4, panel B). This shift is also no longer equivalent to that observed for the C-terminal deletion peptides (compare curves of $TnI_{96-148A2}$, panel B with those of TnI_{96-139} and TnI_{96-131}, Panel A). Truncation of TnI_{96-131} by 16 amino acid residues (116-131) to produce TnI_{96-115} again causes a significant loss in the ability of T•C complex to neutralize the inhibitory activity of this peptide (Fig. III-4, panel A), a result similar to that observed in the TnC release assays (Fig. III-3, panel A). These results continue to suggest (as described above) that residues important for TnC interaction and release of inhibition have now been deleted. Oxidation of the methionine residues (116 and 121 being located within this region) results in the same dramatic loss in the ability of T•C complex to release the inhibitory activity of this peptide (Fig. III-4, panel B, $TnI_{96-148A1}$). Correspondingly, this same effect is also seen when residues 116-126 are substituted with glycines (Fig. III-4, Panel B, $TnI_{96-148A7}$). When the hydrophobic residues (methionine 121, leucine 122 and leucine 125, leucine 126) are substituted with the less hydrophobic residue alanine, the T•C complex release curves are again shifted to higher concentrations and a decreased rate of release is observed, indicating that both sets of hydrophobic residues are important for TnC and T·C complex interaction (Fig. III-4, panel B, $TnI_{96-148A5}$ and $TnI_{96-148A6}$).

Thus, it would appear from these results and those described above (Fig. III-3) that all peptides which contain residues 116-131 intact are able to be quickly neutralized, while analogs which either have this region deleted, modified or substituted are impaired. Moreover, this interaction appears to be a direct result of the interaction of TnC with this region of TnI as it is independent of the presence of TnT.

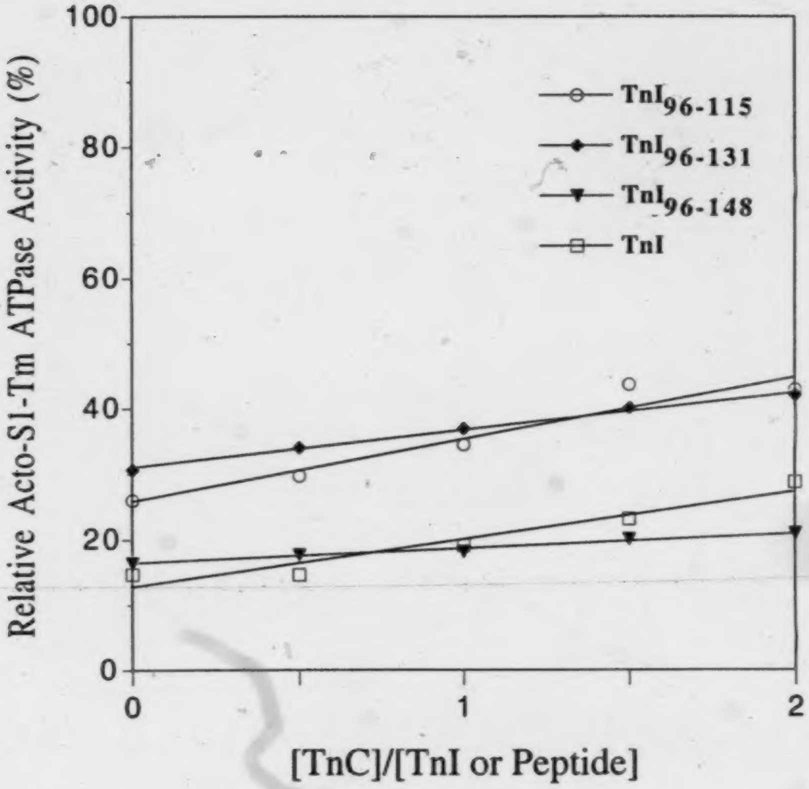


Figure III-5. Neutralization of TnI or TnI peptide inhibition by TnC (minus Ca^{2+}). The ability of TnC (in the absence of $CaCl_2$) to neutralize the inhibitory effect of TnI or TnI peptides was measured. Assays were carried out in a similar manner to that described for TnC (plus Ca^{2+}) assay but in the absence of calcium and in the presence of 80 µM EGTA (pCa ~8.0). All inhibitors were present at a 2:1 mole ratio of inhibitor to Tm. Symbols: TnI (□); TnI_{96-148} (▼); TnI_{96-131} (◆); TnI_{96-115} (O).

Neutralization of the Inhibitory Activity of TnI and TnI Peptides by TnC (minus Ca^{2+}).

In order to determine whether the interaction between TnI residues 116-131 and TnC is Ca^{2+}-dependent, we analyzed the ability of the TnI peptides analogs TnI_{96-115}, TnI_{96-131}, and TnI_{96-148} as well as TnI protein to be released by TnC in the absence of calcium. Results of these assays are shown in Fig. III-5. The addition of TnC (in the absence of calcium) has little effect on the release of inhibition for all of the TnI peptides or protein compared to that observed in the presence of calcium (compare the release at a 2:1 mole ratio; Fig. III-3 and Fig. III-5). Furthermore, the release of the TnI peptides TnI_{96-131} and TnI_{96-148} which contain residues 116-131 show similar rates of release to that of TnI_{96-115}, indicating that the inclusion of residues 116-131 does not increase the ability of these peptides to be released, and thus the significant difference observed in the presence of Ca^{2+} must be a result of a Ca^{2+}-dependent interaction with TnC.

Affinity Chromatography of the TnI Peptides

Taken together, the results of Figs. III-3, III-4 and III-5 indicate that residues 116-131 of TnI are forming an important Ca^{2+}-dependent interaction with TnC. To characterize this interaction further, we determined the relative binding affinities between several of the TnI peptides for TnC using a TnC affinity column. These results are summarized in Table III-2. In the presence of 5 mM $CaCl_2$, TnI_{96-115} peptide, containing only the inhibitory region, bound to the TnC affinity column under starting conditions but was completely eluted by the second gradient step at 300 mM NaCl. In contrast, $TnI_{115-131}$, which bound under starting conditions, was predominantly eluted by the first gradient step (150 mM NaCl) with the remaining 25% of the peptide eluted by the next gradient step at 300 mM NaCl. When both of these regions are combined as in the TnI peptide 96-131, an apparent synergistic binding effect now occurs where the TnI peptide (TnI_{96-131}) can no longer be eluted from the column using NaCl (measured up to 1M). The peptide was in fact finally eluted from the column using a 6 M GdnHCl wash step. Similar results were also seen for the TnI peptide, TnI_{96-148}, which also contains both of these regions. In contrast to the two previous peptides, $TnI_{96-148A1}$, which contains the three oxidized methionine residues, could be eluted with ~500 mM NaCl. The higher salt concentration required to elute this peptide compared to TnI_{96-115} suggests that the "second TnC binding site" has been significantly decreased although not totally disrupted. When the TnI peptide containing the substituted glycine residues within the region 116-126 was analyzed ($TnI_{96-148A7}$), this peptide showed an identical elution profile to that of TnI_{96-115}. This peptide could be completely eluted with 300 mM NaCl, indicating that the second TnC binding site has been totally disrupted. The fact that the "glycine spacer" peptide ($TnI_{96-148A7}$) is eluted in a similar manner to that of TnI_{96-115} also indicates that residues located on the C-terminal side of residue 131 (which are present in $TnI_{96-148A7}$) are apparently not involved in any further stabilizing interaction between this region of TnI and TnC. In addition, this conclusion is further supported by the similar conditions required for elution (6M Gdn•HCl) of TnI_{96-131} and TnI_{96-148} peptides, as well as the observation that residues on the C-terminal side of residue 131 do not appear to be involved in TnC binding as observed in the TnC and T·C release assays (Fig. III-3 and III-4) which only showed dramatic differences in neutralization when residues 116-131 were deleted (or substituted).

Since both oxidation of methionine residues and substitution with the glycine residues is predominantly affecting the hydrophobic residues located within the 116-131 region of TnI, this would suggest that the synergistic binding effect seen for TnI_{96-131} (as compared to the individual peptides TnI_{96-115} and $TnI_{115-131}$) results from two types of binding interactions: (1) a predominantly electrostatic interaction which can be disrupted with salt; and (2) a predominantly hydrophobic interaction which, when combined with the first site is probably stabilized by the higher salt concentrations (and thus GdnHCl is required for peptide elution from the affinity column). Thus, only when the hydrophobic interactions

Table III-2. Affinity Chromatography

TnI Peptide	Elution Conditions[a]				
	NaCl Concentration				
	150 mM	300 mM	500 mM	1000 mM	6 M GdnHCl
Presence of 5 mM Ca^{2+}					
TnI 96-115	—	+++	—	—	—
TnI 115-131	++	+	—	—	—
TnI 96-131	—	—	—	—	+++
TnI 96-148	—	—	—	—	+++
TnI 96-148A1	—	+	++	—	—
TnI 96-148A7	—	+++	—	—	—
Presence of 5 mM Mg^{2+}					
TnI 96-115	+++	—	—	—	—
TnI 115-131[b]	—	—	—	—	—
TnI 96-131	+	++	—	—	—

+++ represents full elution of bound peptide.

++ represents >75% elution of bound peptide.

+ represents <25% elution of bound peptide

a Step gradient conditions: 0 to 15% B in 5 min, followed by a hold for 10 min, followed by an increase to 30% B in 5 min, followed by a hold for 10 min, followed by an increase to 50% B in 5 min, followed by a hold for 10 min, followed by an increase to 100% B in 5 min, followed by a final 10 min hold, where Buffer A is 5 mM aqueous TrisHCl, pH 7, containing 50 mM NaCl and 5 mM $CaCl_2$ or 5 $MgCl_2$, and Buffer B is Buffer A but with 1 M NaCl.

b Not retained by the affinity column.

are removed via oxidation or elimination of their side chains, are the peptides eluted as if they contain only the single 96-115 site.

To verify further that the interactions between residues 116-131 are calcium specific, we repeated the same experiment but in the presence of 5 mM $MgCl_2$ (Table III-2, bottom). TnI_{96-115}, which was retained by the affinity column under starting conditions, was fully eluted by the first gradient step (150 mM NaCl). In contrast, $TnI_{115-131}$ was not retained by the affinity column. When both regions are combined TnI_{96-131} showed only a slight increase in affinity (eluted by the 300 mM NaCl). This is in stark contrast to that observed in the presence of Ca^{2+}, where TnI_{96-131} could not be eluted even at salt concentrations of 1 M. The significant binding affinity differences observed in the presence or absence of calcium agrees with the previous results (Fig. III-5) that the interaction between TnC and the TnI 116-131 region is Ca^{2+}-dependent.

Ca^{2+} -Sensitivity.

We next determined what effect each one of the defined binding regions would have on the Ca^{2+}-sensitivity of release. Fig. III-6, panels A and B show the Ca^{2+}-dependent neutralization by TnC of the inhibitory activity of the three TnI peptides (TnI_{96-115}; TnI_{96-131}; TnI_{96-148}) and TnI as a function of calcium concentration (pCa). Measurements were made essentially the same as in Fig. III-3 except that the free Ca^{2+} concentration was varied at a fixed ratio of TnI peptide:Tm (2:1, respectively). The three TnI peptides tested showed essentially the same pCa_{50} midpoint transition values (6.6 +/- 0.05; Table III-3), although they exhibited quite different Ca^{2+}-dependence release curves. As can be seen, the inhibition of TnI_{96-148} and TnI_{96-131} can be fully released similar to that of intact TnI, while TnI_{96-115} can only be partially released (under these conditions) (Fig. III-6, panel A). When the same experiment was repeated in the presence of TnT (T·C complex, Fig. III-6, panel B) similar Ca^{2+}-dependent release curves were observed with the exception of

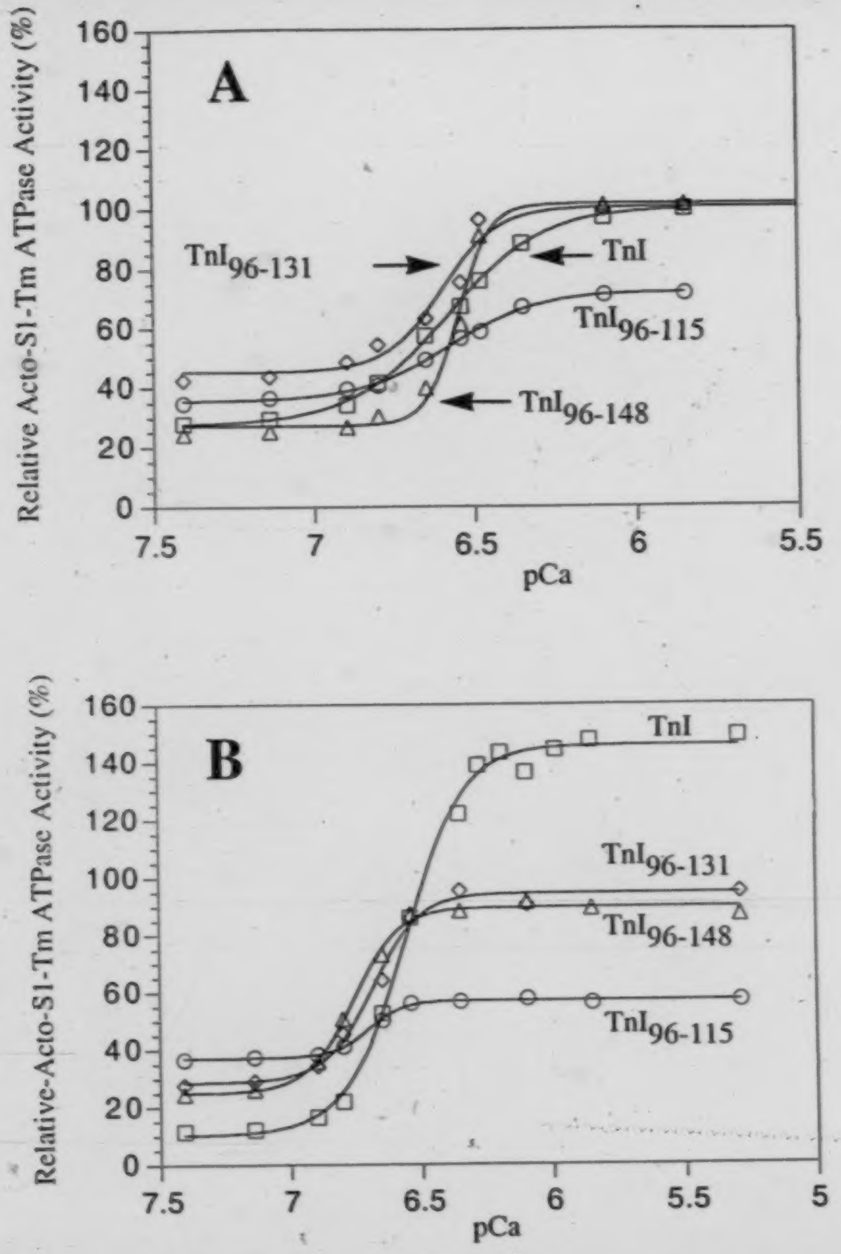


Figure III-6. Ca^{2+}-dependent release of TnI peptides 96-115, 96-131, 96-148 and TnI. The Ca^{2+}-dependence was carried out as described in "Materials and Methods". The three TnI peptides: TnI_{96-115} (O); TnI_{96-131} (◊); TnI_{96-148} (Δ) and TnI (❑) were present at a 2:1 mole ratio to Tm. Panel A, Ca^{2+}-dependent neutralization by TnC. Panel B, Ca^{2+}-dependent neutralization by T·C. Data were fitted to a sigmoidal equation. The pCa_{50} values are shown in Table 2. ATPase activity is expressed as a percentage of the difference between the velocity at pCa 5.5 and pCa 7.5. Each point is an average of 3 determinations.

Table III-3. Titration of TnC and T·C complex with calcium

Peptide/Protein	TnC $pCa_{50}{}^{a,b}$	T·C $pCa_{50}{}^{a,b}$
TnI_{96-115}	6.55	6.70
TnI_{96-131}	6.61	6.68
TnI_{96-148}	6.54	6.76
TnI	6.64	6.65

[a] pCa_{50} is the pCa (-log of the Ca^{2+} concentration) required to induce half of the Ca^{2+}-dependent change in acto-S1-Tm ATPase activity. pCa_{50} mid-points were determined by curve fitting the change in ATPase activity vs. Ca^{2+} concentration for each peptide or protein.

[b] assays were carried out at a mole ratio of 7:2:4:8 of actin:Tm:inhibitor:TnC or T·C in the presence 0.25 mM EGTA and 13 mM KCl.

intact TnI, which activated the ATPase activity (the acto-S1-Tm ATPase activity is taken as 100%).

These results indicate that, although the different binding sites can affect the amount of inhibition or release, they do not appear to affect the Ca^{2+}-dependent transition midpoints (6.7 +/- 0.05; Table III-3), and thus their Ca^{2+} sensitivity. Moreover, these results suggest that the critical region defining the Ca^{2+}-dependent transition midpoint is the inhibitory region (residues 96-115).

Competition Between the N- and C-terminal Regions of TnI

Previous studies have shown that, in addition to the inhibitory region, residues located at the N-terminus of TnI (residues 1-47) can also bind to TnC (Syska *et al.*, 1976). Subsequent investigations into this region have now shown that this region binds to the C-domain of TnC and can block the neutralization of TnI peptide 104-115 or TnI by TnC (Farah *et al.*, 1994, Ngai & Hodges, 1992). It is from these studies that it has been suggested that the N-terminal (residues 1-40; Rp_{1-40}) and C-terminal (residues 96-115) regions of TnI may compete for similar binding sites on the C-domain of TnC. With the above data supplying further evidence that the inhibitory region binds to the C-domain of TnC, we re-investigated the ability of the N- and C-terminal regions of TnI to bind competitively to TnC and whether the binding between these two regions is affected by the presence of residues 116-131 of TnI. To carry out this competition, we used a mole ratio of 4:2:1 TnC:TnI (or inhibitory peptide):Tm. Under these conditions, the inhibition of the inhibitory peptides or TnI are fully released (with the exception of TnI_{96-115}, which was 90% released; Fig. III-3, panel A), but if completely displaced from TnC could produce maximal inhibition (Fig. III-1). We then titrated in Rp_{1-40} up to a ratio of 2:1 (Rp_{1-40}:TnC). Results of the assay are shown in Fig. III-7. Titration of Rp_{1-40} was able to displace TnI_{96-115} from TnC significantly in the presence of Ca^{2+}. All of the other TnI peptides and TnI which include the "second TnC binding site" apparently have a greater

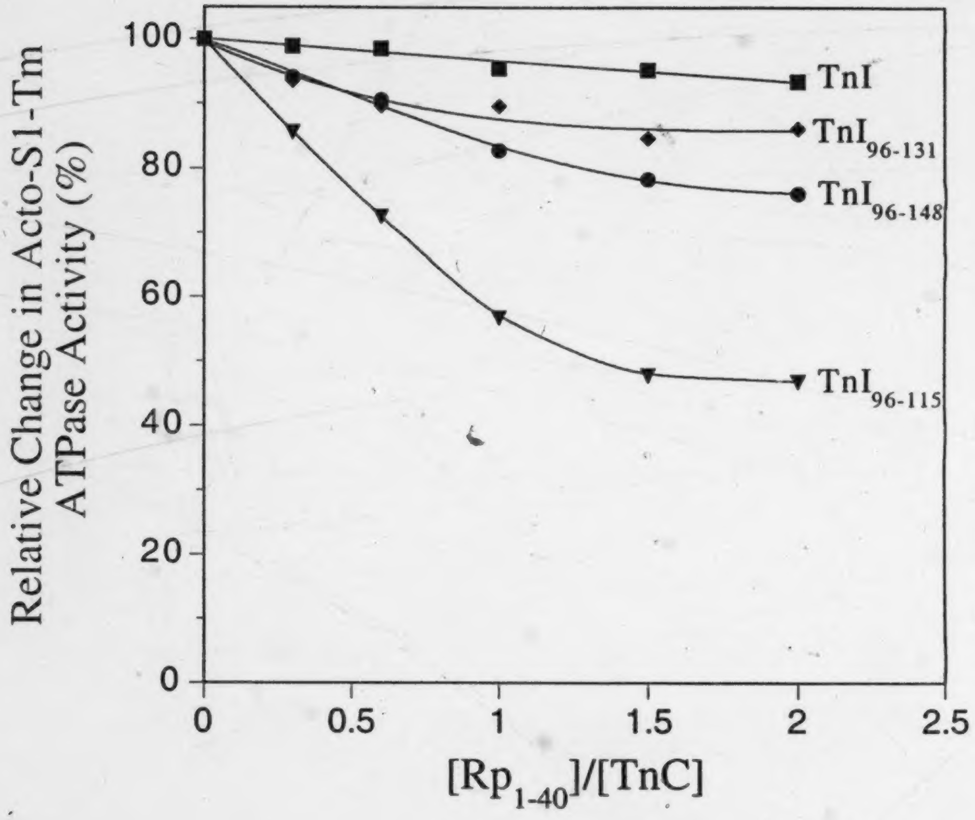


Figure III-7. Competitive displacement of TnI or TnI peptides from TnC (plus Ca^{2+}) by Rp_{1-40}. The effect of Rp_{1-40} on the release of TnI or TnI peptides from TnC (plus Ca^{2+}) was measured by the decrease in ATPase activity with increasing concentrations of Rp_{1-40}. The assay was carried out as described in "Materials and Methods". TnI (■); TnI_{96-131} (◆); TnI_{96-148} (●); TnI_{96-115} (▼). 100% is equal to the acto-S1-Tm ATPase rate in the absence of Rp_{1-40}.

affinity than Rp_{1-40} for TnC, and thus can not be as effectively displaced. Comparison of the displacement of TnI_{96-131} and TnI_{96-148} shows that TnI_{96-148} has a slightly greater decrease in ATPase activity compared to that of TnI_{96-131}. This difference might, however, be explained by the fact that TnI_{96-148} when released is a much better inhibitor than TnI_{96-131} (at equivalent concentrations) due to the presence of the second actin-Tm binding site (Fig. III-1, Panel A). Therefore, the two peptides may in fact be released to similar degrees and only differ due to their inhibitory activity. This slight difference thus illuminates an important point, that the affinity differences between the N- and C-terminal regions of TnI for TnC are also affected by the affinity of the inhibitory region for the thin filament. The observation that TnI was not able to be displaced from TnC to any significant degree was quite surprising. In fact, at first it appeared that these results differed from those previously reported from our own laboratory by Ngai and Hodges (Ngai & Hodges, 1992), in which Rp_{1-40} was able to prevent TnI from binding TnC. The difference between these two experiments, however, may in fact be a result of differences between these two assays, i.e., kinetic vs thermodynamic. In the assay reported by Ngai and Hodges (1992), Rp_{1-40} was first pre-mixed with TnC to form a stable complex prior to its addition to solutions of actin-Tm-TnI (or peptide). In the present assay, Rp_{1-40} was added into a pre-mixed solution of actin:Tm:TnI (or peptide):TnC and allowed to incubate only 10 min before being assayed. Thus, in the present assay we are attempting to measure the effectiveness of Rp_{1-40} in displacing TnI or TnI peptides which have bound to TnC in a Ca^{2+} state; in contrast Ngai and Hodges (1992) were essentially measuring the opposite (the displacement of Rp_{1-40} by TnI or TnI peptides). In this case, the more physiologically relevant of the two assays is unknown, since the first assay may be showing that Rp_{1-40} is able to prevent the inhibitory region from binding TnC; whereas in the second assay, once the complete binding region of TnI_{96-131} is bound to TnC (in the presence of Ca^{2+}), the affinity of the inhibitory region and the "second TnC binding site" now exceeds the affinity of Rp_{1-40}. Both of these results may be correct, and only highlights the actual complexity of the Ca^{2+}-dependent regulatory process. In our earlier work (Ngai & Hodges, 1992), it was mentioned how the presence of the remaining polypeptide chain of TnI is probably critical in regulating how the N-terminal and C-terminal regions of TnI bind to TnC.

Overall, the results suggest that the inhibitory region and the N-terminus of TnI may in fact be binding to similar sites within the C-terminus of TnC. Furthermore, TnI residues 116-131 appear to modulate the effectiveness of this displacement. These results, in conjunction with those described above showing significant differences in affinity, support the Ca^{2+}-dependent switching mechanism between the N- and C-terminal regions of TnI for the C-domain of TnC proposed by Ngai and Hodges (1992).

C. Discussion

Using several synthetic peptides corresponding to various regions of the C-terminus of TnI, we have been able to map out and identify two other physiologically important binding sites of TnI for TnC and actin-Tm. Both of these sites appear to be critical in allowing this region of TnI (residues 96-148) to obtain an inhibitory activity similar to that of intact TnI protein as well as an ability to be released effectively by TnC in a Ca^{2+}-dependent manner. Results of the inhibition and actin-Tm binding studies have been able to identify that both the N-and C-terminal regions (residues 96-115, and ~140-148) of the C-terminus of TnI are involved in binding to the thin filament, as well as mediating an inhibitory activity. Surprisingly, although the region of 96-148 was chosen as a starting point in the mapping of the critical regions required for full inhibition, it appears to represent the minimal sequence of TnI required to achieve this "wild-type like" inhibitory activity. Truncation of the C-terminal 9 residues (140-148) or substitution of the three C-terminal lysine residues (141, 144, and 145) completely abolishes the enhanced inhibitory activity when compared to TnI_{96-115}. Moreover, identification that the central region (116-131) is not important for binding to the actin-TM filament nor the inhibitory activity indicates that the two inhibitory regions (96-115 and 140-148) are non-contiguous.

It is interesting to note the possibility that these same C-terminal residues (~140-148) of TnI could be involved in a second non-contiguous interaction site with actin-Tm was previously suggested by Farah and Reinach (1995) based on the weak homology between residues within the inhibitory region (108-115) and residues 137-144.

```
residues 108-115:  RPPLRRVR
                   |  |* |*
residues 137-144:  RANLKQVK
```

Although this region does appear to show a weak homology to the inhibitory region, this study has been able to demonstrate that it does not appear to be able to induce either an inhibitory activity nor binding affinity similar to that of the inhibitory region (96-115). These results would agree with previous studies which have shown that all the residues within the inhibitory region 104-115 are critical for inducing the inhibitory response and binding affinity to the thin filament (Van Eyk & Hodges, 1988). The fact that there are similarities between these two regions, however, does suggest that either two similar binding sites exist on actin-Tm, or that the inhibitory region and residues ~135-148 span more than one actin monomer. It is interesting to note that, in the binding studies, the TnI peptides containing the inhibitory region did not saturate at a 1:1 mole ratio of actin to peptide. Since the inhibitory region appears to be the dominant binding site, it may be binding to both of these sites. It is also noteworthy that although there was greater than 1:1 binding, this did not exceed a 2:1 molar ratio even at higher concentrations measured. In an attempt to clarify this point further, we are presently attempting to crosslink both of these regions (96-115 and 140-148) simultaneously to the actin-Tm filament to investigate these two possibilities.

Results of the inhibition assays have also demonstrated that the central residues (~116-131) not only play a minimal role in the inhibitory activity, but can in fact antagonize the inhibitory activity in the absence of the C-terminal 9 residues 140-148. These observations thus allude to the interesting possibility that residues 116-131 of TnI do not actually bind to the surface of the thin filament, but rather loop out from the thin filament being anchored only by the inhibitory region and residues ~140-148 on either side. This idea becomes even more attractive when one considers that these same residues (116-131) have now been identified as being involved in a Ca^{2+}-dependent interaction with the N-domain of TnC (data not shown). Thus, it appears that in order for TnI to be quickly released from the thin filament, it requires an exposed site not bound to the thin filament which can be quickly "latched" onto by TnC.

In addition to the C-terminal residues (140-148), this study has also been able to demonstrate clearly the importance of the 8 N-terminal residues (96-103) for the inhibitory activity. Although previous studies have shown that the inhibitory activity does decrease upon the removal of these residues, the present study has been able to ascribe this decrease directly to a loss in the ability of this region to bind the actin filament compared to that of the complete inhibitory region (96-115). It is interesting to note that these same residues of TnI (96-103) have also been shown to enhance the binding interactions between the inhibitory region and TnC (Chandra *et al.*, 1994). Both of these observations now raise new questions regarding the possible role these residues (96-103) may play in "modulating" the inhibitory activity of this region or its release, particularly when residues 96-100 of rabbit fast skeletal muscle (NQKLF), rabbit slow muscle (KLKVL) and rabbit cardiac muscle (TQKIF) differ , and differences in their function have been noted (Guo *et al.*, 1994, Syska *et al.*, 1974, Talbot & Hodges, 1981b).

The Importance of Residues 116-131 for TnC and T·C Neutralization.

Both the TnC and T·C release assays have demonstrated that residues 116-131, located directly adjacent to the inhibitory region, plays a critical role in the Ca^{2+}-dependent interaction with TnC required for the full and rapid neutralization of the inhibitory activity of these peptides. Substitution, oxidation or truncation of residues within this region either completely abolishes or significantly decreases this interaction. The fact that these modifications and substitutions are predominantly affecting only hydrophobic residues within this region demonstrates that the interaction of this region of TnI with TnC is predominantly of a hydrophobic nature. Subsequent studies using affinity chromatography have demonstrated that the presence of this region of TnI can dramatically increase the binding affinity of the inhibitory region for TnC, and thus the significant differences observed in the release assays can be ascribed to a change in the affinity for the C-terminal region of TnI for TnC. Since the residues which bind into the regulatory domain of TnC have not been definitively defined, we therefore propose that it is residues 116-131 of TnI that bind the N-domain of TnC (plus Ca^{2+}) and thus are the critical residues involved in the Ca^{2+}-dependent change in interactions between TnI and TnC.

Further support for the above proposal can also be generated from recent binding, fluorescence and photochemical crosslinking studies between binary complexes of these two proteins or their fragments. For example, Farah *et al*: (1994) showed that the first 102 residues of TnI can be deleted, resulting in the loss of the Ca^{2+}-independent sites of interaction while the Ca^{2+}-dependent interaction sites were not lost. Kobayashi *et al.* (1991, 1994), using TnC mutants with single cysteine residues modified with 4-maleimidobenzophenone at positions 12 and 57, identified that position 57 of rabbit skeletal TnC predominantly crosslinks to TnI residues 113-121, and TnC position 12 to residues 132-141 of TnI, both close to our 116-131 region. Moreover, Tao *et al* (1989, 1990) have shown that TnI modified at cysteine 133 with 4-maleimidobenzophenone can be crosslinked to actin in the absence of calcium, further supporting the idea that residues 140-148 are bound to the actin filament, but move away from the thin filament towards cysteine 98 of TnC in the calcium saturated state. Furthermore, fluorescence studies by Pearlstone and Smillie have clearly shown that residues 117-148 can significantly increase the binding affinity of the inhibitory region for the N-domain of TnC compared to just the inhibitory region (TnI_{96-115})(Pearlstone & Smillie, 1995a, 1997). In addition, the binding orientation of these residues within the hydrophobic pocket of the N-domain of TnC continues to agree with the anti-parallel orientation of the two proteins previously suggested by Farah *et al.* (1994).

Binding of the Inhibitory Region to the Central Helix and C-domain.

The proposal that residues 116-131 of TnI binds the N-domain of TnC further supports the binding of the inhibitory region along the central helix and into the C-domain (Cachia *et al.*, 1983, 1986, Campbell *et al.*, 1991, Campbell & Sykes, 1991, Chandra *et al.*, 1994, Dalgarno *et al.*, 1982, Drabikowski *et al.*, 1985, Grabarek *et al.*, 1981, Howarth *et al.*, 1995, Jha *et al.*, 1996, Kobayashi *et al.*, 1994, 1996, Krudy *et al.*, 1994, Lan *et al.*, 1989, Leszyk *et al.*, 1987, 1988, Ngai *et al.*, 1994, Slupsky *et al.*, 1992, Swenson & Fredricksen, 1992, Tsuda *et al.*, 1992, Wang *et al.*, 1990, Weeks & Perry, 1978). Grabarek *et al.*, (1981), using a variety of proteolytic fragments of TnC, have indicated that the region comprising residues 89-100 of the central helix constitutes one of the two Ca^{2+}-dependent sites of interaction between TnI and TnC based on the observations that only TnC fragments containing these residues are able to regulate the actomyosin ATPase activity when complexed with TnI and TnT, presumably by binding to the inhibitory region of TnI in the presence of calcium. Photochemical crosslinking studies from cysteine 98 of TnC, which is located within the central helix, predominantly labels TnI in the region 103-113 (Leszyk *et al.*, 1987). Similarly, position 89 of TnC can also be crosslinked to this same region (Kobayashi *et al.*, 1994, 1996, Leszyk *et al.*, 1990). In addition, attachment of the photoaffinity probe (benzophenone moiety) to the α-amino group of TnI inhibitory peptide 104-115 results in the labeling of the H-helix (residues 154-159) of the C-domain of TnC (Ngai *et al.*, 1994). Fluorescence binding studies between the inhibitory region and TnC have shown that although the inhibitory region can bind both domains, it binds 20-40 fold stronger to the C-domain of TnC, suggesting the central helix and the C-domain are its preferred binding site (Chandra *et al.*, 1994, Lan *et al.*, 1989, Pearlstone & Smillie, 1995b). Furthermore, we have recently modeled the region 96-131 of TnI bound to the N-

and C-domains of TnC, and observe excellent interactions between the hydrophobic region (residues 116-131) of TnI and the N-domain of TnC , and the inhibitory region (residues 96-115) with the central helix and C-domain of TnC (data not shown).

Structural/Functional Role of the N-terminus of TnI ($Rp_{1\text{-}40}$).

One of the issues which has remained unclear in regards to the interactions between TnI and TnC has come from the identification that the N-terminal residues of TnI also bind to the C-domain of TnC (Farah *et al.*, 1994). Recent studies have shown that the N-terminal region of TnI can bind to the C-domain in both the presence or absence of calcium (plus $MgCl_2$) and is critical for forming and maintaining the presence of TnC in the ternary troponin complex at low calcium concentrations (Potter *et al.*, 1995, Sheng *et al.*, 1992). Furthermore, Ngai and Hodges (1992) have shown that synthetic peptides containing residues 1-40, 10-40, 20-40 can effectively compete with TnI or the TnI inhibitory peptide 104-115 for binding to TnC and block their neutralization by TnC. From these observations, Ngai and Hodges alluded to the possibility that both of these regions (1-40 and 96-115) may bind to similar sites within the C-terminus but are regulated by other Ca^{2+}-dependent changes in the interactions between TnI and TnC. In an attempt to clarify this issue further, particularly in light of the identification of residues 116-131 (the second Ca^{2+}-dependent interaction site), we determined the ability of the N-terminal residues 1-40 to displace the C-terminal inhibitory region from TnC in the presence or absence of these residues . The results clearly demonstrate that the presence of the C-terminal residues 116-131 of TnI can affect the affinity between the N- and C-terminal regions of TnI for TnC. Moreover, the results indicate that, in the absence of the C-terminal residues 116-131, the N-terminal region of TnI dominates in the affinity for TnC (as judged by the quick displacement of the inhibitory peptide, $TnI_{96\text{-}115}$); however, in the presence of these residues, the affinity is apparently reversed (as judged by the limited release of the C-terminal peptides $TnI_{96\text{-}131}$ and $TnI_{96\text{-}148}$). In addition, the slow release of the TnI peptides, TnI_{96-131} and TnI_{96-148} correlates with the affinity chromatography results, which indicates that these peptides have a significantly greater affinity for TnC than those without the "second TnC binding site".

It is interesting to note that, in the displacement of the inhibitory region peptide (TnI_{96-115}), the curve leveled off at a 1:1 ratio of Rp_{1-40}:TnC, suggesting that the N-terminal peptides can only bind a single site within TnC. This is in contrast to that of the inhibitory region which Pearlstone and Smillie (1995b) and Swenson and Fredrickson (1992) have shown can bind to both domains, although with an order of magnitude less strongly to the N-domain. This may in fact explain why the displacement, although rapid initially, leveled off at 50%, i.e., some release is still able to occur through the binding of the inhibitory region to the N-domain of TnC. Overall, however, these observations support the switching mechanism proposed by Ngai and Hodges (1992).

Transmission of the Calcium Signal to TnT-Tm.

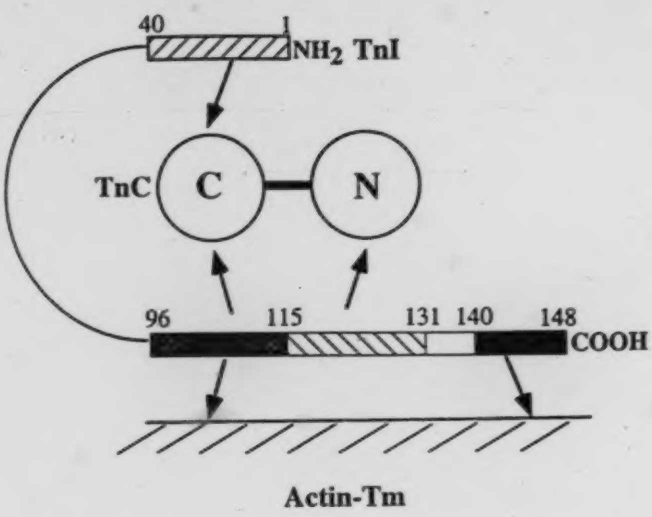
The inability of any of the synthetic TnI peptide analogs in the region 96-148 to activate the acto-S1-Tm ATPase activity in the presence of T·C agrees with earlier results reported by Van Eyk *et al.* (1997). Furthermore, these observations continue to support the idea that the residues important for this activation are located within the N-terminus of TnI (~40-96) (Farah *et al.*, 1994, Potter *et al.*, 1995). The fact that it has now been shown that a TnI fragment of residues 1-98 can activate the acto-S1-Tm ATPase activity even in the absence of the inhibitory region (Farah *et al.*, 1994) suggests that the inhibitory region does not induce the activating conformation between the N-terminal residues of TnI and TnT-Tm but rather only regulates their position or structure. From the above results, we can now propose two possible mechanisms by which the binding of calcium to TnC and its change in interactions with the inhibitory region may be involved in regulating this activating state. In the first mechanism, the binding of the inhibitory region to the actin-Tm thin filament disrupts the interactions of residues ~40-96 of TnI with TnT, ultimately disrupting the activating state. Upon binding calcium, TnC switches the inhibitory region from actin to the C-domain of TnC releasing the disruptive effect and now allowing re-establishment of the interactions of residues 40-96 of TnI with TnT and, hence, re-activation. In the second mechanism, no disruption actually occurs between TnI-TnT interactions located at the amino terminus, but rather the conformational state that can activate the ATPase activity constitutively exists but is prevented from activating by the interactions of the inhibitory region bound to the actin-Tm filament; thus, when the inhibitory region is bound, it holds the activated state in an alternative position on the thin filament (i.e., out of the actin groove). Once this interaction has been released by the binding of TnC in the presence of calcium, the otherwise favorable activating position of TnI-TnT-Tm is assumed (into the groove). In other words, the activation of the ATPase is regulated directly by the movement of the inhibitory region back and forth between actin and TnC with no conformational changes actually transmitted to TnT.

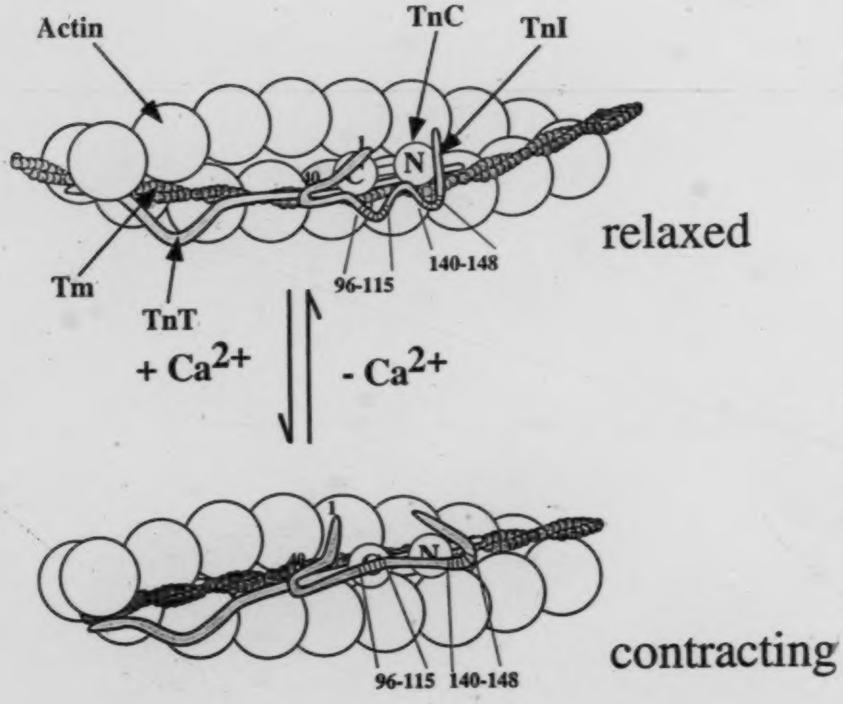
A Model for the Ca^{2+}-dependent Regulation of Muscle Contraction

The mapping of the two distinct binding regions within the C-terminus of TnI cited in the present paper can now be incorporated into a general model describing the Ca^{2+}-induced structural changes that regulate muscle contraction. In this model (Fig. III-8), in the relaxed (Mg^{2+}/low Ca^{2+}) state, TnI and TnC are arranged in an anti-parallel orientation with the N-terminal region of TnI (residues 1-40) bound structurally to the COOH-domain of TnC ((Farah *et al.*, 1994, Farah & Reinach, 1995) and references therein). The inhibitory region of TnI, as well as the second actin-Tm binding region (residues ~140-148), are bound to the actin-Tm filament causing full inhibition of the ATPase activity and muscle relaxation. Upon stimulatory release of calcium, Ca^{2+} ions bind the two low affinity sites (regulatory sites I and II) within the NH_2-domain of TnC, to cause a re-orientation of the N, B and C helices (relative to A and D) and thus expose a hydrophobic pocket (Gagne *et al.*, 1994, 1995, Herzberg *et al.*, 1986). The opening and exposure of this pocket now allows

---

Figure III-8. Schematic summary of the interactions of TnI residues 1-40 and 96-148 with TnC and actin-Tm, and a model depicting their role in the Ca^{2+}-dependent regulation of muscle contraction. Top; The various regions of TnI involved in binding TnC and the actin-Tm are indicated by arrows. Regions of TnI proposed to interact with the N- and C-domain of TnC are hatched. Regions of TnI that bind the actin-Tm thin filament and are responsible for the full inhibitory activity are shaded. Bottom; A schematic model depicting the possible organization and changes occurring in the troponin complex in the absence and presence of Ca^{2+} based on these results and those of others (see "Discussion). N and C refer to the N and C-terminal regions of TnC. TnI, TnT and Tm are indicated in the figure. Circular spheres represent actin monomers. $-Ca^{2+}$ state (relaxed state). In the absence of Ca^{2+}, the N-terminal region of TnI is anchored strongly onto the C-domain of TnC, stabilizing the ternary complex with TnT. The inhibitory region (residue 96-115) and the C-terminal region (140-148) (hatched) of TnI are bound to the actin-Tm filament inhibiting the ATPase activity. The central region (residue 116-131) is looped out from the thin filament. $+Ca^{2+}$ state (contracting state). In the presence of Ca^{2+}, the regulatory sites in the N-terminal domain of TnC bind Ca^{2+} and cause the opening of the hydrophobic pocket. The central residues 116-131 of TnI now bind the newly exposed "pocket" in the N-domain and switch the inhibitory region from the actin-Tm filament to the C-domain of TnC. The movement of the inhibitory region into the C-domain of TnC and displacement of the N-terminal region out of the C-domain of TnC causes a conformational change in residues 40-96 of TnI which indirectly (or directly) affects interactions with TnT and Tm which triggers the activation of the ATPase activity.


40
1
NH2 TnI
TnC
C
N
96
115
131
140
148
COOH
Actin-Tm


Actin
TnC
TnI
1
40
C
N
relaxed
Tm
TnT
96-115
140-148
+ Ca2+
- Ca2+
1
40
C
N
contracting
96-115
140-148

residues 116-131 of TnI, located directly adjacent to the inhibitory region, to bind the N-domain of TnC, and in doing so significantly increases the affinity between TnI and TnC, causing the inhibitory region to switch from the actin-Tm filament to the COOH-domain of TnC thus releasing the inhibitory effect of TnI on the thin filament and initiating the subsequent movement of TnT and Tm which leads to muscle contraction. In our present model, binding of the inhibitory region to the COOH-domain of TnC would displace the NH_2-terminal region of TnI (residues 1-40). Although the NH_2-terminal region of TnI is not directly required for Ca^{2+}-dependent regulation (Farah *et al.*, 1994, Sheng *et al.*, 1992), its movement off the COOH-domain of TnC is proposed to work in concert with the inhibitory region to re-activate the activation of the ATPase activity which is observed only in the presence of TnT (Farah *et al.*, 1994, Potter *et al.*, 1995, Sheng *et al.*, 1992). As indicated in the model (Fig. III-8), the region of TnI which interacts with TnT (and is now implicated in this effect) is comprised of residues ~40-96, located between these two sites. It has been proposed that this region of TnI may form a two-stranded coiled-coil with TnT (Pearlstone & Smillie, 1985); thus the possible formation or breakage of this region may initiate TnT movement.

When the signal for contraction has ceased, and the Ca^{2+} concentration decreased, the loss of Ca^{2+} ions in the regulatory sites of the N-domain of TnC causes the N-domain to release residues 116-131 of TnI and close its hydrophobic pocket. Loss of this N-terminal interaction drastically reduces the affinity of the TnI for TnC such that the inhibitory region can now be competitively displaced from the COOH-domain of TnC by the NH_2-terminal region of TnI (residues 1-40), thus returning the TnI-TnC-TnT interactions of the troponin complex back to the low calcium state, and relaxation.

D. Conclusion

In summary, the above study was undertaken to investigate the functional role of the C-terminal residues 116-148 of TnI in the Ca^{2+} dependent regulation of skeletal muscle contraction. The above results indicate that residues 116-131, located directly adjacent to the inhibitory region, bind to the N-domain of TnC in a Ca^{2+}-dependent manner and are critical for the Ca^{2+}-dependent neutralization of the inhibitory region (96-115). Residues ~140-148 bind the actin-Tm filament and appear to be critical for allowing the 96-148 region of TnI to achieve an inhibitory activity similar to that of intact TnI. Although none of the inhibitory peptides analyzed are able to activate the ATPase activity in the presence of TnT, these results only further support the proposal that the residues important for this function are located within the N-terminus of TnI (residues 40-96). Lastly, the series of competition assays between the N and C terminal regions of TnI have now shown that the presence of the "second TnC binding site" (residues 116-131) can significantly affect the relative binding affinity strengths between these two regions of TnI (1-40 and 96-115) for TnC, and thus may both bind to the C-domain of TnC but are regulated by calcium and the binding of this region (residues 116-131).

Overall, these results re-confirm and expand the important role that TnI plays as a Ca^{2+} dependent molecular switch in thin filament-based regulation of muscle contraction.

E. Related Research

Since reporting this work, the interactions between the synthetic peptide $TnI_{115-131}$ and the Ca^{2+} saturated N-domain of TnC have been studied using NMR spectroscopy. These results show the $TnI_{115-131}$ binding site on N-TnC is in the exposed "hydrophobic pocket". Moreover, utilizing the change in amide chemical shifts, the dissociation constant was determined to be 24 μM with a off rate constants of 300 s^{-1}. These results are reported in the paper entitled "Interaction of the second binding region of troponin I with the regulatory domain of skeletal muscle troponin C as determined by NMR spectroscopy" by R.T McKay, B.Tripet, R.S. Hodges and B.D. Sykes published in *J. Biol. Chem* 1997, 272, 28494-28500.

F. References

Cachia, P. J., Sykes, B. D., and Hodges, R. S. (1983) Calcium-Dependent Inhibitory Region of Troponin: A Proton Nuclear Magnetic Resonance Study on the Interaction between Troponin C and the Sythetic Peptide Nα-Acetyl[FPhe^{106}]TnI-(104-115) Amide. *Biochemistry* **22,** 4145-4152.

Cachia, P. J., Van Eyk, J., Ingraham, R. H., McCubbin, W. D., Kay, C. M., and Hodges, R. S. (1986) Calmodulin and troponin C: a comparative study of the interaction of mastoparan and troponin I inhibitory peptide [104-115]. *Biochemistry* **25,** 3553-3562.

Campbell, A. P., Cachia, P. J., and Sykes, B. D. (1991) Interaction of troponin I and troponin C: 19F NMR studies of the binding of the inhibitory troponin I peptide to turkey skeletal troponin C. *Biochem. & Cell Biol.* **69,** 674-681.

Campbell, A. P. & Sykes, B. D. (1991) Interaction of troponin I and troponin C. Use of the two-dimensional nuclear magnetic resonance transferred nuclear Overhauser effect to determine the structure of the inhibitory troponin I peptide when bound to skeletal troponin C. *J. Mol. Biol.* **222,** 405-421.

Chandra, M., McCubbin, W. D., Oikawa, K., Kay, C. M., and Smillie, L. B. (1994) Ca^{2+}, Mg^{2+}, and troponin I inhibitory peptide binding to a Phe-154 to Trp mutant of chicken skeletal muscle troponin C. *Biochemistry* **33,** 2961-2969.

Chong, P. C. S., Asselbergs, P. J., and Hodges, R. S. (1983) Inhibition of rabbit skeletal muscle acto-S1 ATPase by troponin T. *FEBS Lett.* **153,** 372-376.

Chong, P. C. S. & Hodges, R. S. (1982) Photochemical Cross-linking between Rabbit Skeletal Troponin Subunits. *J. Biol. Chem.* **257,** 11667-11672.

Collins, J. H., Potter, J. D., Horn, M. J., Wilshire, G., and Jackman, N. (1973) The Amino Acid Sequence Of Rabbit Skeletal Muscle Troponin C: Gene Replication and Homology with Calcium-Binding Proteins from Carp and Hake Muscle. *FEBS Lett.* **36,** 268-272.

Dalgarno, D. C., Grand, R. J. A., Levine, B. A., Moir, A. J. G., Scot, G. M. M., and Perry, S. V. (1982) Interaction between troponin I and troponin C. *FEBS Lett.* **150,** 54-59.

Drabikowski, W., Dalgarno, D. C., Levine, B. A., Gergely, J., Grabarek, Z., and Leavis, P. C. (1985) Solution conformation of the C-terminal domain of skeletal troponin C. *Eur. J. Biochem.* **151,** 17-28.

Ebashi, S. & Endo, M. (1968) Calcium ion and muscle contraction. *Prog. Biophys. Mol. Biol.* **18,** 123-183.

Ebashi, S., Wakabayashi, T., and Ebashi, F. (1971) Troponin and Its Components. *J. Biochem.* **69,** 441-445.

Farah, C. S., Miyamoto, C. A., Ramos, C. H., da Silva, A. C., Quaggio, R. B., Fujimori, K., Smillie, L. B., and Reinach, F. C. (1994) Structural and regulatory functions of the NH2- and COOH-terminal regions of skeletal muscle troponin I. *J. Biol. Chem.* **269,** 5230-5240.

Farah, C. S. & Reinach, F. C. (1995) The troponin complex and regulation of muscle contraction. [Review]. *FASEB J.* **9,** 755-767.

Fujimori, K., Sorenson, M., Herzberg, O., Moult, J., and Reinach, F. C. (1990) Probing the calcium-induced conformational transition of troponin C with site-directed mutants. *Nature* **345,** 182-184.

Gagne, S. M., Tsuda, S., Li, M. X., Chandra, M., Smillie, L. B., and Sykes, B. D. (1994) Quantification of the calcium-induced secondary structural changes in the regulatory domain of troponin-C. *Prot. Sci.* **3,** 1961-1974.

Gagne, S. M., Tsuda, S., Li, M. X., Smillie, L. B., and Sykes, B. D. (1995) Structures of the troponin C regulatory domains in the apo and calcium-saturated states. *Nature Structural Biology* **2,** 784-789.

Grabarek, Z., Drabikowski, W., Leavis, P. C., Rosenfeld, S. S., and Gerggely, J. (1981) Proteolytic Fragments of Troponin C. *J. Biol. Chem.* **256,** 13121-13127.

Grabarek, Z., Tan, R. Y., Wang, J., Tao, T., and Gergely, J. (1990) Inhibition of mutant troponin C activity by an intra-domain disulphide bond [see comments]. *Nature* **345,** 132-135.

Grabarek, Z., Tao, T., and Gergely, J. (1992) Molecular mechanism of troponin-C function. [Review]. *J. Mus. Res. & Cell Motility* **13,** 383-393.

Greaser, M. L. & Gergely, J. (1971) Reconstitution of troponin activity from three protein components. *J. Biol. Chem.* **246,** 4226-4233.

Greaser, M. L. & Gergely, J. (1973) Purification and properties of the components from troponin. *J. Biol. Chem.* **248,** 2125-2133.

Guo, X., Wattanapermpool, J., Palmiter, K. A., Murphy, A. M., and Solaro, R. J. (1994) Mutagenesis of cardiac troponin I. Role of the unique NH_2-terminal peptide in myofilament activation. *J. Biol. Chem.* **269,** 15210-15216.

Heinonen, J. K. & Lahti, R. J. (1981) A New and Convenient Determination of Inorganic Orthophosphate and its Application to the Assay of Inorganic Pyrophosphatase. *Anal. Biochem.* **113,** 313-317.

Herzberg, O. & James, M. N. G. (1985) Structure of the calcium regulatory muscle protein troponin-C at 2.8 A resolution. *Nature* **313,** 653-659.

Herzberg, O., Moult, J., and James, M. N. (1986) A model for the Ca^{2+}-induced conformational transition of troponin C. A trigger for muscle contraction. *J. Biol. Chem.* **261,** 2638-2644.

Hichcock, S. E. (1981) Study of the Structure of Troponin-C by Measuring the Relative Reactivities of Lysines with Acetic Anhydride. *J. Mol. Biol.* **147,** 153-173.

Howarth, J. W., Krudy, G. A., Lin, X., Putkey, J. A., and Rosevear, P. R. (1995) An NMR and spin label study of the effects of binding calcium and troponin I inhibitory peptide to cardiac troponin C. *Prot. Sci.* **4,** 671-680.

Ingraham, R. H. & Hodges, R. S. (1988) Effects of Ca^{2+} and subunit interactions on surface accessibility of cysteine residues in cardiac troponin. *Biochemistry* **27,** 5891-5898.

Jha, P. K., Mao, C., and Sarkar, S. (1996) Photo-cross-linking of rabbit skeletal troponin I deletion mutants with troponin C and its thiol mutants: the inhibitory region enhances binding of troponin I fragments to troponin C. *Biochemistry* **35,** 11026-11035.

Kobayashi, T., Leavis, P. C., and Collins, J. H. (1996) Interaction of a troponin I inhibitory peptide with both domains of troponin C. *Biochimica et Biophysica Acta* **1294,** 25-30.

Kobayashi, T., Tao, T., Gergely, J., and Collins, J. H. (1994) Structure of the troponin complex. Implications of photocross-linking of troponin I to troponin C thiol mutants. *J. Biol. Chem.* **269,** 5725-5729.

Kobayashi, T., Tao, T., Grabarek, Z., Gergely, J., and Collins, J. H. (1991) Cross-linking of residue 57 in the regulatory domain of a mutant rabbit skeletal muscle troponin C to the inhibitory region of troponin I. *J. Biol. Chem.* **266,** 13746-13751.

Kretsinger, R. H. & Nockolds, C. E. (1973) Carp muscle-binding protein II. Structure determination and general description. *J. Biol. Chem.* **248,** 3313-3326.

Krudy, G. A., Kleerekoper, Q., Guo, X., Howarth, J. W., Solaro, R. J., and Rosevear, P. R. (1994) NMR studies delineating spatial relationships within the cardiac troponin I-troponin C complex. *J. Biol. Chem.* **269,** 23731-23735.

Lan, J., Albaugh, S., and Steiner, R. F. (1989) Interactions of troponin I and its inhibitory fragment (residues 104-115) with troponin C and calmodulin. *Biochemistry* **28,** 7380-7385.

Leszyk, J., Collins, J. H., Leavis, P. C., and Tao, T. (1987) Cross-linking of rabbit skeletal muscle troponin with the photoactive reagent 4-maleimidobenzophenone: identification of residues in troponin I that are close to cysteine-98 of troponin C. *Biochemistry* **26,** 7042-7047.

Leszyk, J., Collins, J. H., Leavis, P. C., and Tao, T. (1988) Cross-linking of rabbit skeletal muscle troponin subunits: labeling of cysteine-98 of troponin C with 4-maleimidobenzophenone and analysis of products formed in the binary complex with troponin T and the ternary complex with troponins I and T. *Biochemistry* **27,** 6983-6987.

Leszyk, J., Grabarek, Z., Gergely, J., and Collins, J. H. (1990) Characterization of zero-length cross-links between rabbit skeletal muscle troponin C and troponin I: evidence for direct interaction between the inhibitory region of troponin I and the NH2-terminal, regulatory domain of troponin C. *Biochemistry* **29,** 299-304.

Ngai, S. M. & Hodges, R. S. (1992) Biologically important interactions between synthetic peptides of the N-terminal region of troponin I and troponin C. *J. Biol. Chem.* **267,** 15715-15720.

Ngai, S. M., Sonnichsen, F. D., and Hodges, R. S. (1994) Photochemical cross-linking between native rabbit skeletal troponin C and benzoylbenzoyl-troponin I inhibitory peptide, residues 104-115. *J. Biol. Chem.* **269,** 2165-2172.

Pato, M. O., Mak, A. S., and Smillie, L. B. (1981) Fragments of Rabbit Striated Muscle α-Tropomyosin. *J. Biol. Chem.* **256,** 593-601.

Pearlstone, J. R. & Smillie, L. B. (1985) The interaction of rabbit skeletal muscle troponin-T fragment with troponin-I. *Can. J. Biochem. Cell. Biol.* **63,** 212-218.

Pearlstone, J. R. & Smillie, L. B. (1995a) Ca^{2+}-Specific Fluorescence Changes Upon Interaction of Troponin-I Inhibitory Peptides with the N and C Domains of Troponin-C Tryptophan Mutants. *Biophys. J.* **68,** A166.

Pearlstone, J. R. & Smillie, L. B. (1995b) Evidence for two-site binding of troponin I inhibitory peptides to the N and C domains of troponin C. *Biochemistry* **34,** 6932-6940.

Pearlstone, J. R. & Smillie, L. B. (1997) Interactions of Structural C Domain and Regulatory N Domain of Troponin C with Reapeated Sequence Motif in Troponin I. *Biophys. J.* **72,** A331.

Perry, S. V., Cole, H. A., Head, J. F., and Wilson, F. J. (1972) Localization and mode of action of the inhibitory protein component of the troponin complex. *Cold Spring Harbor Symp. Quant. Biol.* **37,** 251-262.

Potter, J. D., Sheng, Z., Pan, B. S., and Zhao, J. (1995) A direct regulatory role for troponin T and a dual role for troponin C in the Ca^{2+} regulation of muscle contraction. *J. Biol. Chem.* **270,** 2557-2562.

Schechter, Y. (1986) Selective oxidation and reduction of methionine residues in peptides and proteins by oxygen exchange between sulfoxide and sulfide. *J. Biol. Chem.* **261,** 66-70.

Sereda, T. J., Mant, C. T., Sonnichsen, F. D., and Hodges, R. S. (1994) Reversed-phase chromatography of synthetic amphipathic α-helical peptides as a model for ligand/receptor interactions: Effect of changing hydrophobic environment on the relative hydrophilicity/hydrophobicity of amino acid side-chains. *J. Chromatogr.* **676,** 139-153.

Sheng, Z., Pan, B. S., Miller, T. E., and Potter, J. D. (1992) Isolation, expression, and mutation of a rabbit skeletal muscle cDNA clone for troponin I. The role of the NH2 terminus of fast skeletal muscle troponin I in its biological activity [published erratum appears in J Biol Chem 1993 Feb 5;268(4):3016]. *J. Biol. Chem.* **267,** 25407-25413.

Sheng, Z., Strauss, W. L., Francois, J. M., and Potter, J. D. (1990) Evidence that both Ca(2+)-specific sites of skeletal muscle TnC are required for full activity [published erratum appears in J Biol Chem 1993 May 5;268(13):9936]. *J. Biol. Chem.* **265,** 21554-21560.

Slupsky, C. M., Reinach, F. C., Smillie, L. B., and Sykes, B. D. (1995) Solution secondary structure of calcium-saturated troponin C monomer determined by multidimensional heteronuclear NMR spectroscopy. *Prot. Sci.* **4,** 1279-1290.

Slupsky, C. M., Shaw, G. S., Campbell, A. P., and Sykes, B. D. (1992) A 1H NMR study of a ternary peptide complex that mimics the interaction between troponin C and troponin I. *Prot. Sci.* **1,** 1595-1603.

Slupsky, C. M. & Sykes, B. D. (1995) NMR solution structure of calcium-saturated skeletal muscle troponin C. *Biochemistry* **34,** 15953-15964.

Spudich, J. A. & Watts, S. (1971) The regulation of rabbit skeletal muscle contraction. 1. Biochemical studies of the interaction of the tropomyosin-troponin complex with actin and the proteolytic fragments of myosin. *J. Biol. Chem.* **246,** 4866-4871.

Sundaralingam, M., Bergstrom, R., Strasburg, G., Rao, S. T., Roychowdhury, P., Greaser, M., and Wang, B. C. (1985) Molecular Structure of Troponin C from Chicken Skeletal Muscle at 3-Angstrom Resolution. *Science* **227,** 945-948.

Swenson, C. A. & Fredricksen, R. S. (1992) Interaction of troponin C and troponin C fragments with troponin I and the troponin I inhibitory peptide. *Biochemistry* **31,** 3420-3429.

Syska, H., Perry, S. V., and Trayer, I. P. (1974) A New Method of Preparation of Troponin I (Inhibitory Protein) Using Affinity Chromatography. Evidence for Three Different Forms of Troponin I in Striated Muscle. *FEBS Lett.* **40,** 253-257.

Syska, H., Wilkinson, J. M., Grand, R. J. A., and Perry, S. V. (1976) The Relationship between Biological Activity and Primary Structure of Troponin I from White Skeletal Muscle of the Rabbit. *Biochem. J.* **153,** 375-387.

Talbot, J. A. & Hodges, R. S. (1979) Synthesis and Biological Activity of an Icosapeptide Analog of the Actomyosin ATPase Inhibitory Region of Troponin I. *J. Biol. Chem.* **254,** 3720-3723.

Talbot, J. A. & Hodges, R. S. (1981b) Synthetic Studies on the Inhibitory Region of Rabbit Skeletal Troponin I. *J. Biol. Chem.* **256,** 2798-2802.

Tao, T., Gong, B. J., and Leavis, P. C. (1990) Calcium-induced movement of troponin-I relative to actin in skeletal muscle thin filaments. *Science* **247,** 1339-1341.

Tao, T., Gowell, E., Strasburg, G. M., Gergely, J., and Leavis, P. C. (1989) Ca^{2+} dependence of the distance between Cys-98 of troponin C and Cys-133 of troponin I in the ternary troponin complex. Resonance energy transfer measurements. *Biochemistry* **28,** 5902-5908.

Tsuda, S., Aimoto, S., and Hikichi, K. (1992) 1H-NMR study of Ca(2+)-and Mg(2+)-dependent interaction between troponin C and troponin I inhibitory peptide (96-116). *J. Biochem.* **112,** 665-670.

Van Eyk, J. E. & Hodges, R. S. (1988) The biological importance of each amino acid residue of the troponin I inhibitory sequence 104-115 in the interaction with troponin C and tropomyosin-actin. *J. Biol. Chem.* **263,** 1726-1732.

Van Eyk, J. E., Strauss, J. D., Hodges, R. S., and Ruegg, J. C. (1993) A synthetic peptide mimics troponin I function in the calcium-dependent regulation of muscle contraction. *FEBS Lett.* **323,** 223-228.

Van Eyk, J. E., Thomas, L. T., Tripet, B., Wiesner, R. J., Pearlstone, J. R., Farah, C. S., Reinach, F. C., and Hodges, R. S. (1997) Distinct regions of Troponin I regulate Ca^{2+}-dependent activation and Ca^{2+}-sensitivity of the acto-S1-Tm ATPase activity of the thin filament. *J. Biol. Chem.* **272,** 10529-10537.

Wang, Z. Y., Sarkar, S., Gergely, J., and Tao, T. (1990) Ca2(+)-dependent interactions between the C-helix of troponin-C and troponin-I. Photocross-linking and fluorescence studies using a recombinant troponin-C. *J. Biol. Chem.* **265,** 4953-4957.

Weeds, A. G. & Taylor, R. S. (1975) Seperation of subfragment-1 isoenzymes from rabbit skeletal muscle myosin. *Nature* **257,** 54-56.

Weeks, R. A. & Perry, S. V. (1978) Characterization of a region region of the primary sequence of troponin C involved in calcium ion-dependent interaction with troponin I. *Biochem. J* **173,** 449-457.

Yagi, K., Yazawa, Y., and TUsutumu, Y. (1967) Proteolytic Separation of an enzymatic Active Subfragment from the Myosin-Subfragment-S-1. *Biochem. Biophys. Res. Commun.* **29,** 331-336.

Zot, A. S. & Potter, J. D. (1987) Structural aspects of troponin-tropomyosin regulation of skeletal muscle contraction. [Review]. *Annu. Rev. Biophys. & Biophys. Chem.* **16,** 535-559.

CHAPTER IV

Demonstration of Coiled-Coil Interactions within the Kinesin Neck Region using Synthetic Peptides: Implications for Motor Activity

A version of this work has been published: Brian Tripet, Ronald D. Vale and Robert S. Hodges, *J. Biol. Chem.*, 1997, 272,8946-8956.

A. Introduction

Understanding how motor proteins generate force and movement from the chemical hydrolysis of ATP remains one of the most intriguing problems in biophysics. At present, there are three separate families of motor proteins found within eukaryotic cells: myosins, which move along actin filaments, and kinesins and dyneins, which move along microtubules (Vale, 1993). The motor domains that typify each of these superfamilies exhibit little or no amino acid sequence similarity, and hence it was believed that they had evolved separately and were structurally unrelated. However, the recently determined crystal structure of kinesin revealed an unexpected structural similarity to the core of the myosin motor domain, particularly in the nucleotide binding pocket (Kull *et al.*, 1996; Sablin *et al.*, 1996). Hence, myosin and kinesin may share some similarities in how they generate unidirectional movement and force, although the precise mechanistic details remain to be elucidated for both types of motors.

Kinesin has proven to be an excellent model system for investigating the mechanism of motility, in part due to the small size of its motor domain (>2-fold smaller than myosin's). Kinesin purified from tissue sources exists as an $\alpha_2\beta_2$ heterotetramer, in which two α subunits (heavy chains) and two β subunits (light chains) associate to form a highly elongated molecule with globular termini (Bloom *et al.*, 1988; Kuznetsov *et al.*, 1988). The kinesin heavy chains are organized into four domains (listed from N to C-terminus): i) a ~325 amino acid residue globular motor domain head that contains the ATP and microtubule binding sites, ii) a ~50 amino acid residue region adjacent to the globular motor domain (termed the neck region) that is sufficient for allowing dimerization of the motor domains (Huang *et al.*, 1994) and contains a sequence that is predicted to form an α-helical coiled coil (Stewart *et al.*, 1993; Huang *et al.*, 1994), iii) a long (~450 amino acid residue) α-helical coiled-coil domain (termed the stalk), and iv) a small globular C-terminus (termed the tail) (Hirokawa *et al.*, 1989; Scholey *et al.*, 1989; Yang *et al.*, 1989; de Cuevas *et al.*, 1992). Flexible "hinge" regions are found between the neck and the stalk and in the center of the stalk. The light chains (β subunits) of kinesin, which are not necessary for force-generation, are associated with the smaller globular C-terminus of the heavy chain and may be involved in determining cargo specificity (Hirokawa *et al.*, 1989).

A single kinesin molecule can move continuously along a microtubule for several microns in a series of 8 nm steps, which corresponds to the distance between tubulin binding sites along the microtubule protofilament (Svoboda *et al.*, 1993). Such processive movement, which is not displayed by muscle myosin or ciliary dynein, very likely represents a specialized adaptation that enables few kinesin motors to transport membrane organelles efficiently within cells. Functional studies on recombinantly expressed kinesin heavy chains have begun to uncover regions that are necessary for kinesin motility. Bacterial expression of the first 340 amino acids of the Drosophila kinesin heavy chain (which contains the core N-terminal globular motor domain and the first ~10 amino acids of the neck) produces a monomeric protein that generates directed motility when many motors are interacting simultaneously with a single microtubule in gliding motility assays (Yang *et al.*, 1990; Stewart *et al.*, 1993). However, these monomeric kinesins do not exhibit processive movement when assayed as individual motors in a single molecule fluorescence motility assay (Vale *et al.*, 1996) or a bead assay (Berliner et al., 1995). A kinesin motor containing the complete motor and neck domains, on the other hand, forms a dimer and also exhibits processive movement (Hackney, 1995; Vale *et al.*, 1996). Collectively, these studies suggest that the dimeric structure of kinesin is not essential for force-generation per se, although it does appear to be required for processive movement. This raises the possibility that processive movement may involve a hand-over-hand coordination of the two kinesin heads.

The proposed α-helical coiled-coil domain in the kinesin neck may be structurally important for coordinating the activities of the two kinesin heads during processive movement. The existence of a coiled-coil structure in close proximity to the motor domains, however, raises important questions concerning its exact boundaries and stability, since the connection between the heads must be sufficiently extensible to allow the two motor domains to span the distance between two tubulin dimers during movement. It is also important to determine the thermodynamic properties of the neck coiled-coil to ascertain if it could partially or totally "un-coil" during the generation of a power stroke. Unfortunately, the atomic resolution structure of the neck domain is unknown, since the segment between residues 323 to 349 is disordered and hence invisible in the present electron density maps of human kinesin (hK349)(Kull *et al.*, 1996). Thus, to gain insight into the structure of the kinesin neck region and its possible functional roles, we have investigated the secondary structure of the human kinesin heavy chain neck region using several synthetic peptides in conjunction with CD spectroscopy.

In the present study, we report that a two stranded, α−helical coiled-coil dimerization domain exists between residues 330-369 within the human kinesin neck region, as predicted from previous work. Residues located to the COOH-terminus of this region, 370-383, appear to be unstructured and are not significantly involved in further stabilization of the dimerization domain. Residues located to the NH_2-terminus of the proposed coiled-coil dimerization domain may adopt a β-sheet secondary structure. Analysis of the stability of each peptide indicates that the heptads required to form a stable coiled-coil domain are arranged in a strong-weak-strong manner. Loss of two heptads from either the N- or C-terminus significantly affects dimer stability. These results suggest that a conformational change in the motor domain, driven by a free energy change associated with ATP hydrolysis, could be transmitted in a manner that affects the stability and/or conformation of the adjacent neck region. We propose a model for kinesin motility in which unwinding of a portion of the coiled-coil domain plays an important role in the mechanochemical cycle.

B. Results

CD analysis of the kinesin neck peptides

To examine the structural characteristics of the human kinesin neck region, several synthetic peptides (see Fig. IV-1 for sequences) that span various regions between residues 316 to 383 were prepared and analyzed by circular dichroism (CD) spectroscopy. Figure IV-2 panels A and B show the far UV CD spectra for the kinesin peptides K1-5. Kinesin peptides K2, K3, K4 and K5 all show characteristic α−helical spectra with double minima at 208 and 222 nm (Chen *et al.*, 1972; Gans *et al.*, 1991). The K2 peptide, which represents residues 330-369 of the kinesin neck region, displayed the greatest molar ellipticity (-24,600 degrees), corresponding to ~70% α−helical content or ~28 helical residues (see Table IV-1). The addition of two heptads (14 residues) onto the COOH-terminus of this region (K5, 330-383) results in a decrease in the molar ellipticity to -19,500 degrees indicating that the C-terminal residues 370-383 added are not α−helical. The calculation of the helical content for K5 is 54% or ~29 helical residues, which indicates no loss of the existing helical residues from the region 330-369. The addition of two heptads onto the NH_2-terminus of this region (K4, 316-369) also results in a significant decrease in the molar ellipticity to -13,300 degrees, again indicating the addition of non-helical residues

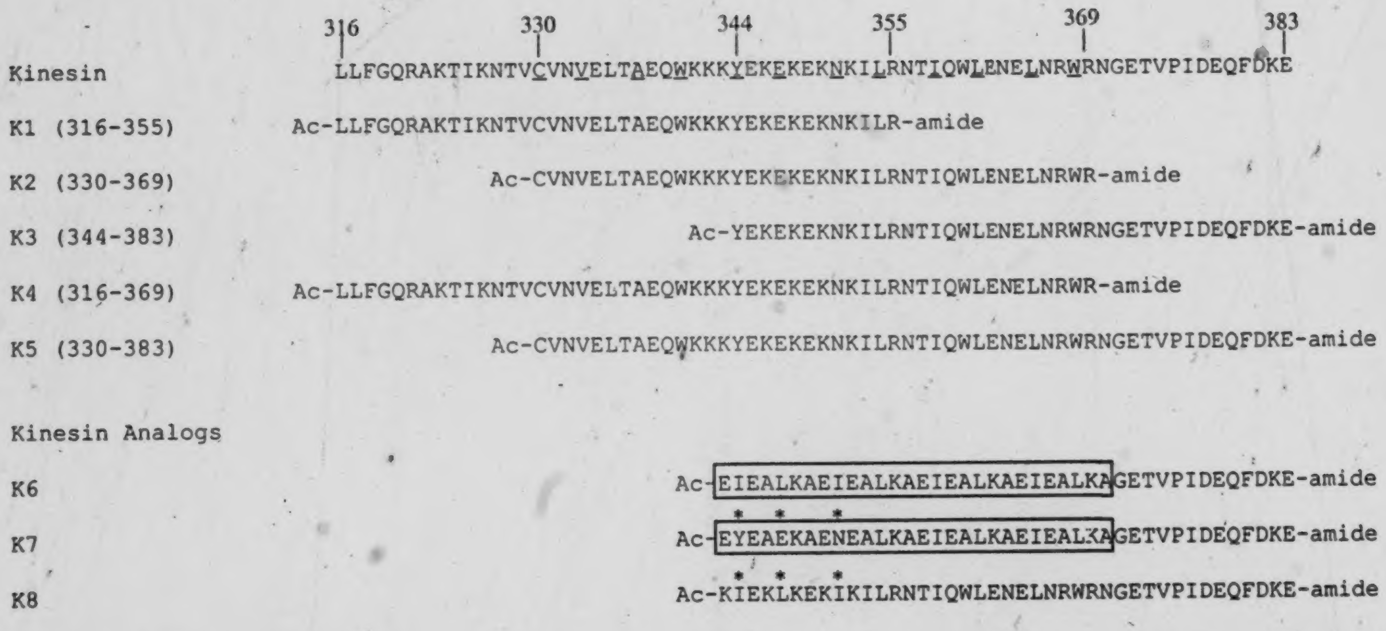

316
330
344
355
369
383
Kinesin
LLFGQRAKTIKNTVCVNVELTAEQWKKKYEKEKEKNKILRNTIQWLENELNRWRNGETVPIDEQFDKE
K1 (316-355)
Ac-LLFGQRAKTIKNTVCVNVELTAEQWKKKYEKEKEKNKILR-amide
K2 (330-369)
Ac-CVNVELTAEQWKKKYEKEKEKNKILRNTIQWLENELNRWR-amide
K3 (344-383)
Ac-YEKEKEKNKILRNTIQWLENELNRWRNGETVPIDEQFDKE-amide
K4 (316-369)
Ac-LLFGQRAKTIKNTVCVNVELTAEQWKKKYEKEKEKNKILRNTIQWLENELNRWR-amide
K5 (330-383)
Ac-CVNVELTAEQWKKKYEKEKEKNKILRNTIQWLENELNRWRNGETVPIDEQFDKE-amide
Kinesin Analogs
K6
Ac-EIEALKAEIEALKAEIEALKAEIEALKAGETVPIDEQFDKE-amide
K7
* * *
Ac-EYEAEKAENEALKAEIEALKAEIEALKAGETVPIDEQFDKE-amide
K8
* * *
Ac-KIEKLKEKIKILRNTIQWLENELNRWRNGETVPIDEQFDKE-amide

Figure IV-1. Amino acid sequences of the kinesin peptides used in this study. The native human kinesin sequence is shown at the top. Residues proposed to be involved in forming the 3-4 hydrophobic repeat of a coiled-coil are underlined. Peptide names are indicated on the left. Numbers in parentheses indicate the region of the amino acid sequence which the peptide spans. Ac- denotes N$^{\alpha}$-acetyl, and -amide denotes C$^{\alpha}$-amide. The three kinesin analogs are shown below. Residues of the native kinesin sequence which have been replaced by a model coiled-coil sequence are boxed. The location of amino acid substitutions in either the native or model coiled-coil sequences are indicated by a star (*). The position of all kinesin peptides and analogs are shown relative to the native sequence.

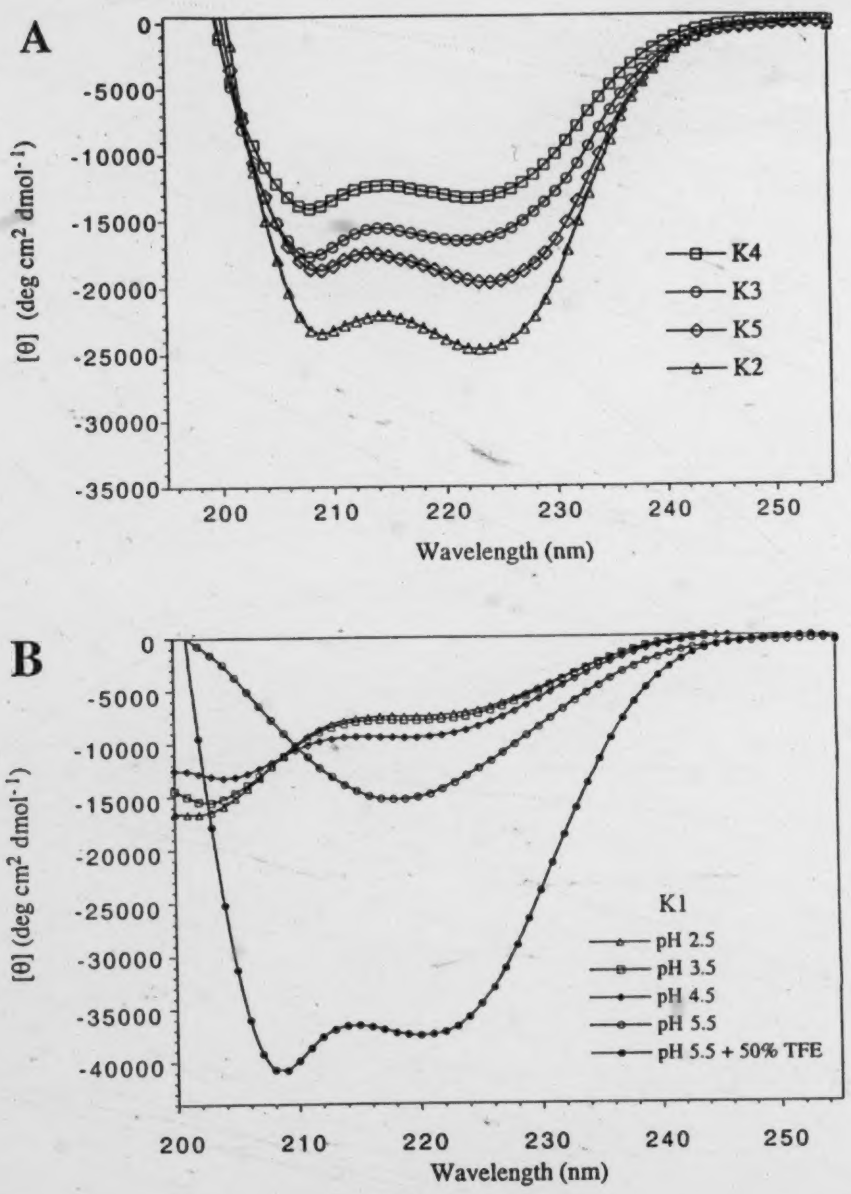


Figure IV-2. Circular Dichroism (CD) spectra of the kinesin peptides. Panel A; CD spectra for kinesin peptides K2-5; Spectra were recorded at 20°C in a 0.1 M KCl, 0.05 M PO_4, 0.002 M DTT, pH 7 buffer. All peptide concentrations were 100 uM. Panel B; the CD spectra of kinesin peptide K1 at various pH's and in the presence of 50% TFE. Spectra were recorded at 20°C in a 0.1 M KCl, 0.05 M PO_4, 0.002 M DTT buffer at the indicated pH. Peptide concentrations were 350 uM.

onto the helical domain located between 330-369. Calculation of the helical content for the peptide K4 is 37% or ~20 helical residues, which represents a loss of ~8 helical residues from the region 330-369. Analysis of the kinesin peptide K3 (344-383), which represents a 40 residue peptide shifted two heptads (14 residues) toward the COOH-terminus, shows a molar ellipticity at 222 nm of -16,500 degrees, corresponding to a helical content of 46% or ~19 helical residues. This represents a loss of ~9 helical residues for the deletion of the two heptads 330-343, indicating that the two NH_2-terminal heptads of 330-343 are very important for the helical content observed in the region 330-369. One can also see from the spectra of the kinesin peptides in Figure IV-2 panel A that the ratios of $[\theta]_{222}$ to $[\theta]_{208}$, which are often used as an indication of coiled-coil formation, are >1 for K2 and K5, but <1 for K3 and K4, suggesting a transition from a possible two-stranded α-helical coiled-coil to a single stranded α-helix (Engel *et al.*, 1991; Lau *et al.*, 1984; Monera *et al.*, 1993; Zhou *et al.*, 1992a,b; Zhu *et al.*,1992).

CD analysis of the kinesin peptide K1 (residues 316-355), which represents a 40 residue peptide shifted two heptads toward the NH_2-terminus (from the region 330-369), reveals a complete absence of α−helical content. In fact the secondary structure of this peptide now displays a β-sheet pattern (Chen *et al.*, 1972). It is important to note, however, that the spectrum of this peptide could not be acquired under the same benign conditions like those used for the other peptides. At pH 2.5, K1 is highly soluble and shows only a random coil spectrum. Increasing the pH successively from 2.5 to 5.5 results in a major ransition from a largely random coil spectrum to that of a β-sheet spectrum, with the greatest transition occurring between pH 4.5-5.5. Analysis of the peptide above pH 5.5 was not possible due to the complete gelation of the solution. The pH dependence of this transition (pH 4.5 to 5.5) suggests that the ionization of glutamic acid residues are involved. Unfortunately, it is not possible to ascertain from these results whether formation of the β-sheet secondary structure is a result of intramolecular or intermolecular interactions, or a combination of both. However, our observations with K4 (which

Table IV-1. Ellipticities and stabilities of the synthetic kinesin peptides

Peptide[a]	$[\theta]_{222}$ (deg·cm^2·dmol^{-1})[b]		Helical Content[c]		$[\mathrm{GdnHCl}]_{1/2}$[d]	$\Delta G_u^{H_2O}$[e]	m[f]	$\Delta\Delta G_u$[g]
	Benign	50%TFE	(%)	#Res.	(M)	(kcal·mol^{-1})		
K2	-24,600	-26,600	70	28	3.61	10.42	1.205	0.0
K2 oxid.	-26,600	-32,200	74	30	5.22	—	—	—
K3	-16,500	-19,190	46	19	1.17	8.46	2.285	1.96
K4	-13,300	-18,440	37	20	3.73	10.14	1.125	0.283
K5	-19,500	-21,200	54	29	3.93	10.31	1.143	0.105
K5 oxid.	-23,200	-27,900	64	34	5.54	—	—	—
K6	-25,000	-25,560	70	29	> 7.0	>20	—	—
K7	-22,200	-25,380	62	25	2.73	10.18	1.208	—
K8	-18,700	-20,500	52	21	3.58	11.20	1.195	2.74*

a The amino acid sequences for each peptide is shown in Fig. 1.

b The mean residue molar ellipticities at 222 nm were measured at 20°C in benign buffer (0.1 M KCl, 0.05 M PO_4, pH 7). For samples containing TFE, the above buffer was diluted 1:1 (v/v) with TFE. Peptide concentrations were 100 μM

c The (%) helical content was calculated from the ratio of the observed $[\theta]_{222}$ value divided by the predicted molar ellipticity x 100. The predicted molar ellipticity was calculated from the equation $[\theta]_{222} = 40 \times 10^3 \times (1\text{-}4.6/N)$ for the chain length dependence of an α-helix (24, 58), where N is the number of residues in the peptide. The number of helical residues was calculated by multipling the % helical content x the total number of residues in the peptide.

d $[\mathrm{GdnHCl}]_{1/2}$ is the transition midpoint, the concentration of guanidine hydrochloride (M) required to give a 50% decrease in the molar ellipticity at 222 nm.

e $\Delta G_u^{H_2O}$ is the free energy of unfolding in the absence of guanidine hydrochloride and was estimated by extrapolating the free energy of unfolding at each denaturant concentration to zero concentration assuming the linear relationship $\Delta G_u = \Delta G^{H_2O} - m[\mathrm{GdnHCl}]$ (18,19). ΔG_u was calculated from the equation $\Delta G_u = -\,RT \ln (2P_t\,(f_u^2/(1\text{-}f_u))$ for reduced peptides. f_u is the molar fraction of denatured peptide as determined from the ellipticity at 222 nm and P_t is the total peptide concentration (M).

f m is the slope in the equation $\Delta G_u = \Delta G_u^{H_2O} - m[\mathrm{GdnHCl}]$.

g The difference in the free enegy of unfolding ($\Delta G_u^{H_2O}$) relative peptide K2.

* Indicates the stability difference relative to peptide K3.

contains this region) showing a loss in helical content of ~8 residues but no substantial loss in stability compared to K2 (discussed below) suggests that their is some intramolecular formation at the N-terminus of this peptide. The gelation of the solution also suggests the formation of intermolecular association as well.

Interestingly, in the presence of 50% TFE, a helix inducing solvent (Lau *et al.*, 1984; Sönnichen *et al., 1992*), the K1 peptide reverts to a fully α-helical spectrum (equal to that calculated for a 40 residue peptide (Chen *et al.*, 1972; Gans *et al.*, 1991)). This result indicates that this region of kinesin has the intrinsic ability to adopt either β-sheet or helical secondary structures depending on the environment. It should be noted that the first few amino acids of this peptide (316-320) are in a helical configuration in the kinesin crystal structure (Kull et al., 1996).

Oxidation of Cys$_{330}$

Although the α-helical content of the K2 peptide (residues 330-369) is significantly greater than that of the other native kinesin peptides, it still does not represent a fully helical structure as calculated theoretically for a 40 residue peptide (~-36,500°) (Gans *et al.*, 1991). We therefore determined if oxidation of the N-terminal cysteine (cys330) to form a disulfide bridge could increase the helical content to that of the theoretical value by stabilizing the ends of the proposed coiled-coil as well as by making the coiled-coil dimerization domain concentration independent (Zhou *et al.*, 1992b, 1992c and 1993). When the peptide was oxidized, a change in its molar ellipticity was observed (see Fig. IV-3). Oxidation only slightly increased the molar ellipticity at 222 nm by approximately 2000 degrees, which is similar to that obtained for the reduced K2 peptide in the presence of 50% TFE, a helix inducing solvent (Lau *et al.*, 1984; Sönnichen *et al.*, 1992). This degree of ellipticity indicates that the reduced (monomeric) K2 peptide is almost fully helical (93%) if judged by the maximal amount of helical content that can be induced either in an oxidized state or in 50% TFE. Although previous studies have shown that theoretical maximum

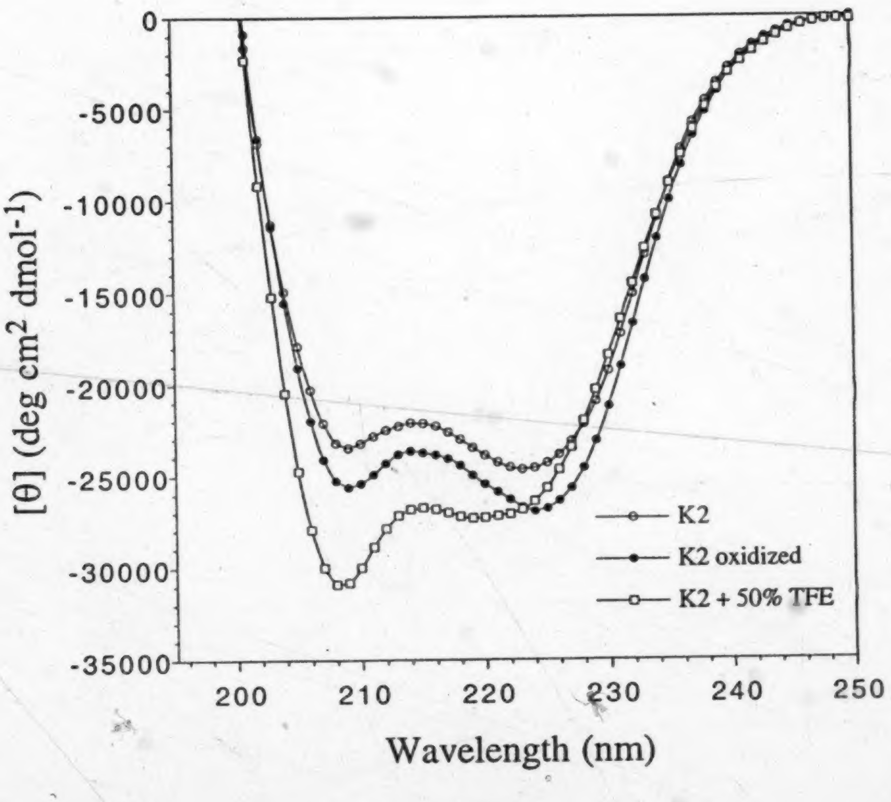


Figure IV-3. CD spectra of the K2 peptide in the reduced and oxidized state, as well as in the presence of 50% TFE. Spectra were recorded in a 0.1 M KCl, 0.05 M PO_4, pH 7 buffer. 2 mM DTT was present for the reduced and 50% TFE scans. Peptide concentrations were 100 uM.

values are not always observed for helical peptides, the observation of a lower molar ellipticity than the theoretical value may also be indicating that there is a region within residues 330-369 that cannot be induced into a fully α-helical structure by either a helix-inducing solvent or oxidation. The central region containing residues Tyr_{344}, Glu_{347} and Asn_{351} in the hydrophobic core is a good candidate for such a region and will be discussed further below.

Two further points can be made regarding the oxidation results. First, the finding that the α–helical secondary structure for the K2 peptide is enhanced instead of disrupted by disulfide bond formation indicates that the two peptides interact in a parallel and in-register manner. Second, while disulfide bond formation of cys330 can occur in a peptide, this may not be possible when the neck is joined to the globular motor domain. This cysteine is also not conserved among conventional kinesin motor proteins from different species.

Stability and concentration dependence of the kinesin peptides

An important question raised regarding the existence of a coiled-coil within the neck region was whether it alone is sufficiently stable to account for the dimerization of the kinesin motor domain heads or whether other subunit-subunit interactions are also involved. In order to address this question, we determined the stability of the five kinesin peptides by GdnHCl denaturation (Fig. IV-4, panel A). The K2 peptide (330-369), which showed the greatest α-helical content (Fig. IV-2), displayed a GdnHCl midpoint of 3.61 M and an extrapolated $\Delta G_u^{H_2O}$ of unfolding of 10.42 kcal/mol, indicating a very stable α-helical structure. The K4 peptide (316-369), which showed a decrease in the helical content and calculated α-helical residues, showed a similar GdnHCl midpoint and $\Delta G_u^{H_2O}$ of unfolding (compare K2 and K4, Table IV-1). Thus the apparent loss of helical residues at the NH_2-terminus of the coiled-coil region is apparently compensated by the formation of an alternative structure or interaction of similar stability.

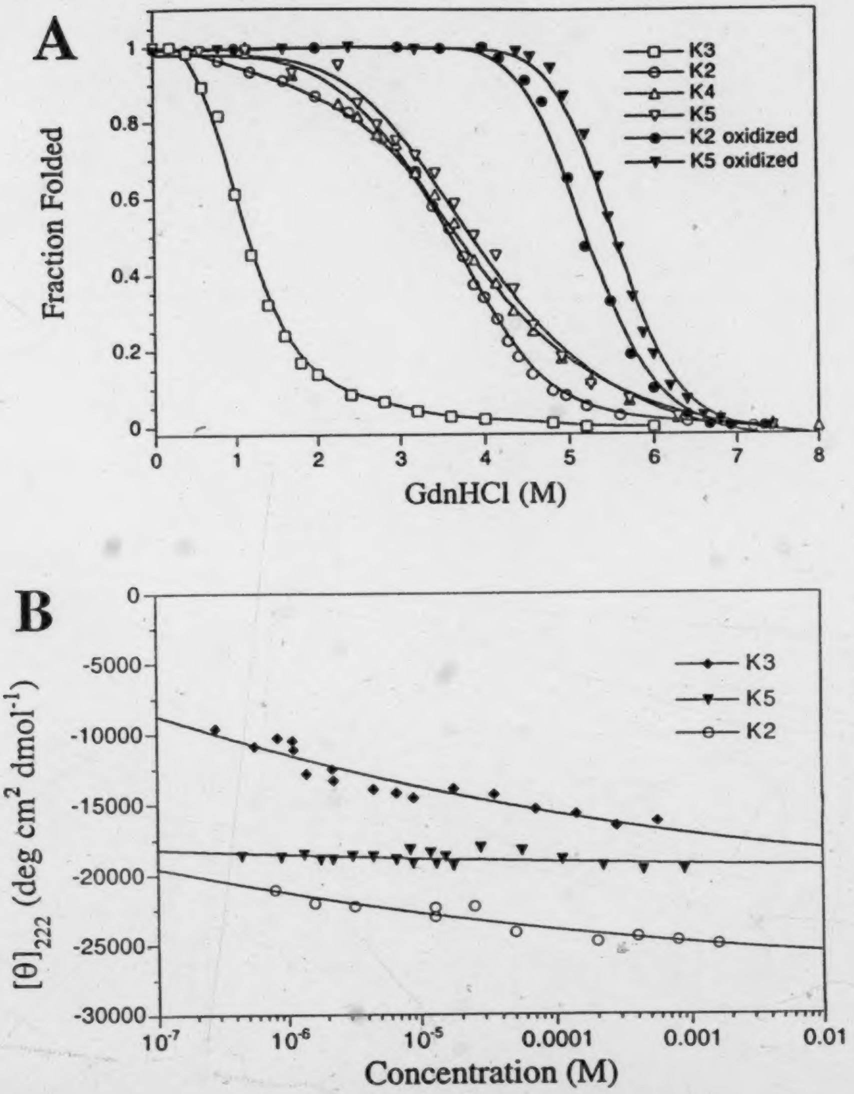


Figure IV-4. Panel A, Denaturation profiles of kinesin peptides K2 - 5 at 20°C in 0.1 M KCl, 0.05 M PO_4, 0.002 M DTT (omitted when oxidized), pH 7 buffer with GdnHCl denaturant. The fraction folded (f_f) of each peptide was calculated as $f_f = ([\theta]-[\theta]_u)/([\theta]_n-[\theta]_u)$, where $[\theta]$ is the observed mean residue ellipticity at 222 nm at any particular denaturant concentration and $[\theta]_n$ and $[\theta]_u$ are the mean residue ellipticities at 222 nm of the native "folded" and "unfolded" states, respectively. Each peptide was analyzed at a 60 uM concentration. Panel B, concentration dependence of the mean residue molar ellipticity at 222 nm for kinesin peptides K2, K3 and K5. Ellipticities were recorded in the same buffer as described above, using various pathlength cells (0.05, 0.1 and 1 cm) depending on the peptide concentration.

The K5 peptide (330-383), the COOH-terminal residues (370-383) of which are non-helical, also showed a similar GdnHCl midpoint and $\Delta G_u^{H_2O}$ of unfolding (compare K2 and K5, Fig. IV-3A, and Table IV-1), indicating that the COOH-terminal residues do not significantly affect the stability of the α-helical structure. The small difference that is observed in the GdnHCl midpoints may be due to end effect stabilization (Zhou *et al.*, 1992c). The stability of the K3 peptide (344-383), which showed a decrease in the helical content and loss of helical residues, was significantly destabilized (GdnHCl midpoint of 1.17 M) by the loss of the two NH_2-terminal heptads (330-343), indicating that these two heptads are very important in the stability of the proposed α-helical coiled-coil structure. Oxidation of both the K2 and K5 peptides resulted in significant increases in their GdnHCl midpoints (3.61 M to 5.22 M and 3.93 M to 5.54 M, respectively), which indicates that the formation of a disulfide bonds stabilizes the α-helical structure as seen in other studies (Zhou *et al.*, 1992b, 1992c and 1993).

The concentration dependence of the α-helical content, which can also be used as an indicator of the stability of the associated coiled-coils, was determined for kinesin peptides K2, K3, and K5. Figure IV-4, panel B, shows that the helical content for K2 and K5 is largely unaffected by concentration over the range tested, indicating that they are very tightly associated α-helical structures. The equilibrium association constants from the GdnHCl denaturation plots for K5 and K2 are estimated to be 2.7×10^8 M^{-1} and 2.3×10^8 M^{-1}, respectively. The similar values for K2 and K5 indicate that the the stability and association of the structure resides principally within residues 330-369. The effect of deleting the two NH_2-terminal heptads (residues 330-343) from K5 (giving peptide K3) shows a greater dependence of the α-helical content with peptide concentration. Quite surprisingly, however, the difference in the concentration dependence of K3 and K2 was not as dramatic as expected from the difference in their stability (Fig. IV-4A). This result may indicate that electrostatic interactions also play a significant role in the association between the two α-helices in K3 (predominantly electrostatic interactions are quickly quenched by GdnHCl denaturation and thus not seen as a major stabilizing factor).

Size exclusion chromatography with laser light scattering

In order to determine the oligomerization state of the α-helical structures in an aqueous solution, size-exclusion chromatography with laser light scattering detection was conducted. Representative SEC chromatograms for the K5 and K3 peptides are shown in Figure IV-5. Analysis of kinesin peptide K5 in the oxidized and reduced states showed only a single eluting peak with an apparent molecular weight obtained from light scattering of 11,702 Da. This value is close to that predicted for a dimeric structure (13,460 Da). The observation of only a single species indicates that K5 peptide forms a very stable, dimeric structure, which agrees with the data presented previously. Similar studies conducted with K2 (data not shown) also produced only one peak corresponding to a dimeric molecular weight. On the other hand, the K3 peptide eluted in two peaks: the first dominant peak having an apparent molecular weight of 9,285 Da and the second smaller peak having a molecular weight (5,854 Da) close to that of the monomeric peptide (5,074 Da). The observation of dimer and monomer peaks is consistent with previous data showing that the K3 peptide exhibits the greatest concentration dependence for helical content. Thus, SEC, CD spectroscopy, and stability studies all suggest that a stable, two-stranded α-helical coiled-coil can form between two chains of the kinesin neck region and that residues 330-369 are the ones of primary importance for the formation of this structure.

Destabilizing effects of Tyr$_{344}$, Glu$_{347}$ and Asn$_{351}$ in the hydrophobic core

Previous studies have shown that the stability of two-stranded α–helical coiled-coils is dependent upon the helical propensity of the region, hydrophobicity of the residues in the core, packing of residues in the core, electrostatic interactions adjacent to the core, and chain length effects (Graddis *et al.* 1993; Kohn *et al.*, 1995a,b; Monera *et al.*, 1993;

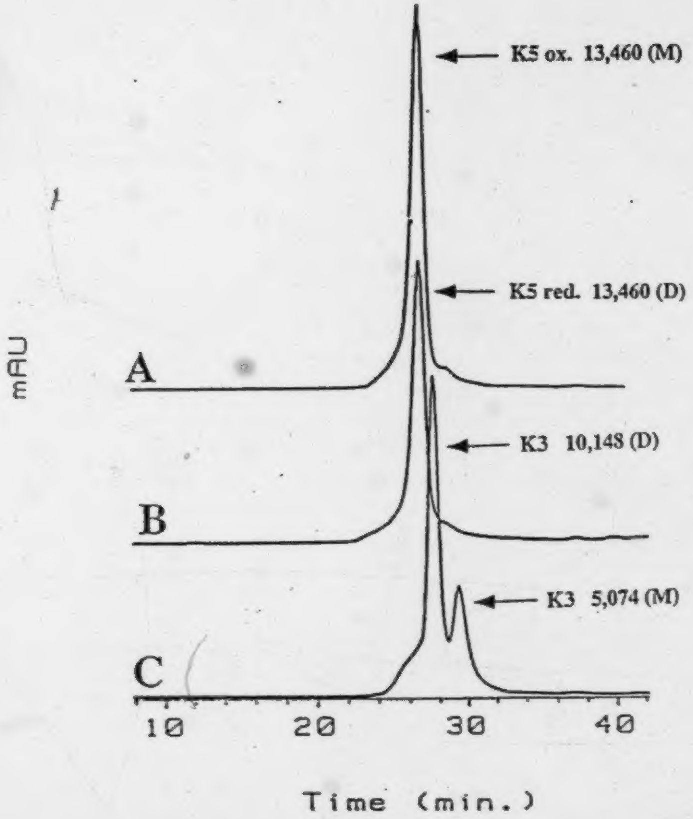


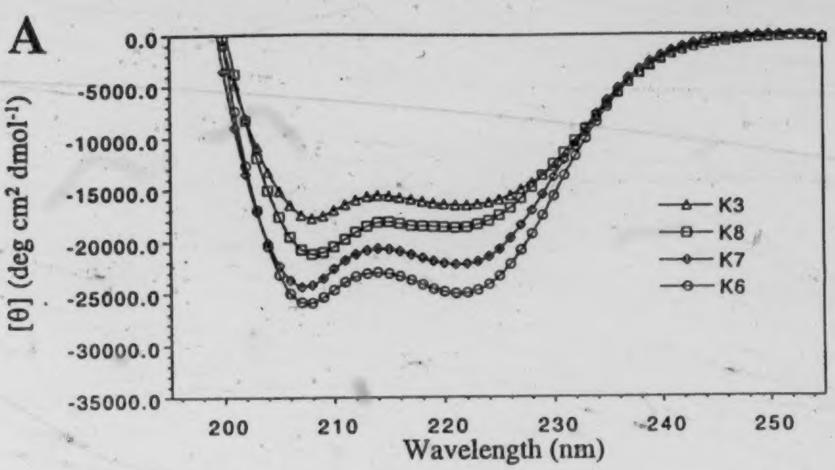
Figure IV-5. Size exclusion chromatography (SEC) of the kinesin peptides. Runs: A) K5, oxidized; B) K5, reduced; C) K3. Column: Superose 12 (1.0 cm x 30.0 cm) from Pharmacia. Running conditions: mobile phase, 100 mM KCl, 50 mM K_2HPO_4, pH 7 buffer; flow rate, 0.5 ml/min.; temperature, 26°C; detection, 210 nm using a Hewlett Packard UV-visible spectrophotometer. Molecular weights for each peptide are indicated as well as the observed oligomerization state: (M) denotes monomer; (D) denotes dimer.

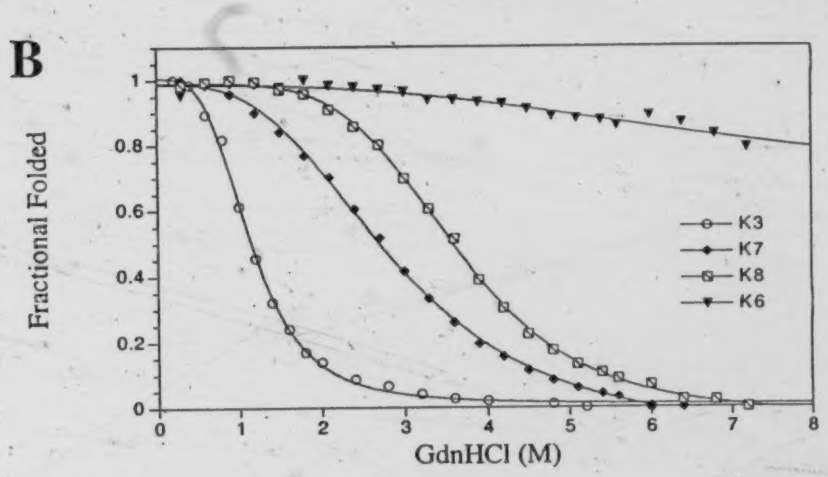
O'Shea *et al.*, 1991; Shuermann *et al.*, 1991; Su *et al.*, *1994*; Zhou *et al.*, 1992a,b; Zhu *et al.*, 1993)). In the kinesin neck region, several of the residues that are predicted to exist within the hydrophobic core are considered to be "non-ideal" for generating a stable coiled-coil. Particularly noticeable are residues Y_{344}, E_{347} and N_{351} that score relatively low by hydrophobicity analysis and thus are not expected to contribute significantly to stability. In particular, the ionized carboxyl group of glutamic acid has been shown to be extremely destabilizing in model coiled-coils (unpublished results). To examine the effects of residues Y_{344}, E_{347}, and N_{351} on coiled-coil stability, [illegible] prepared and analyzed three analog peptides (see Fig. IV-1 for sequences). First, we prepared a kinesin peptide analog (K6) in which the four heptads of the native kinesin sequence between residues 344-370 were replaced by a model coiled-coil sequence that has been previously characterized (Su *et al.*, 1994). Second, in the K7 analog, the three "destabilizing" kinesin hydrophobic core residues, Y_{344}, E_{347} and N_{351}, were substituted into the above model coiled-coil sequence. Finally, in K8, three high stability hydrophobic core residues from the "model" coiled-coil (leucine and isoleucine) were substituted into the native kinesin sequence into positions 344, 347 and 351.

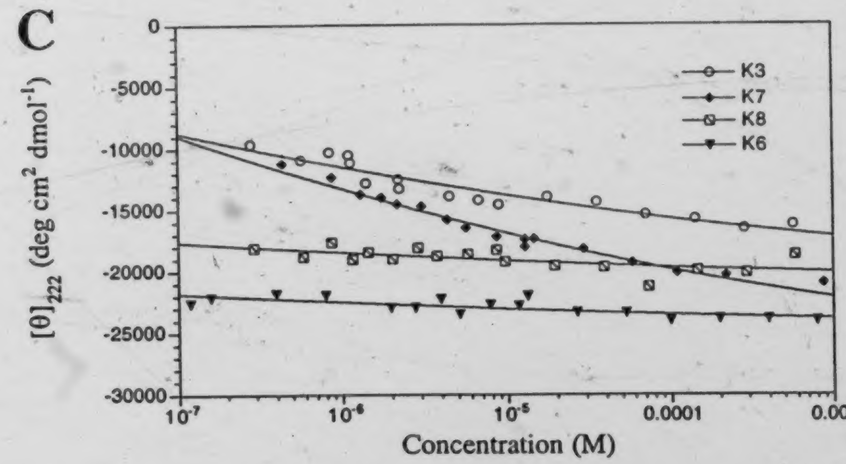
Figure IV-6, panel A shows the secondary structure content of the three kinesin analogs relative to the unsubstituted kinesin peptide K3 (344-383). All three of the kinesin analogs show characteristic double minimas at 208 and 222 nm typical for α-helical protein structures (Chen *et al.*, 1972; Gans *et al.*, 1991). As expected, K6 (the model coiled-coil sequence) displayed the greatest molar ellipticity (Su *et al.*, 1994). Introduction of the three kinesin hydrophobic core residues Y_{344}, E_{347}, and N_{351} into the K6 sequence caused a significant decrease in the molar ellipticity (compare K6 and K7, Fig IV-6A), suggesting a disruption between the two associated α-helices as well as a decrease in coiled-coil stability. Conversely, replacement of Y_{344}, E_{347}, and N_{351} with the "ideal" hydrophobes from the model sequence in K8 resulted in an increase in molar ellipticity,

---

Figure IV-6. Effects of substituting Tyr_{344}, Glu_{347} and Asn_{351}. Panel A, CD spectra for the kinesin analogs K6, K7 and K8 were recorded at 20°C in 0.1 M KCl, 0.05 M PO_4, 0.002 M DTT, pH 7 buffer. Peptide concentrations were 100 uM. Panel B; GdnHCl denaturation profiles of K6, K7 and K8 were recorded at 222 nm at 20°C in 0.1 M KCl, 0.05 M PO_4, 0.002 M DTT, pH 7 buffer. The fraction folded (ff) was calculated as described in Figure IV-4. Peptide concentrations were 60 uM. Panel C, concentration dependence of the mean residue molar ellipticty at 222 nm were measured at 20°C in 0.1 M KCl, 0.05 M PO_4, 0.002 M DTT, pH 7 buffer as described in Figure IV-4. For direct comparison of the three analogs to that of the native sequence, K3 (previously shown in Figures IV-2 and IV-4) is shown again in all three plots.


A
[θ] (deg cm² dmol⁻¹)
0.0
-5000.0
-10000.0
-15000.0
-20000.0
-25000.0
-30000.0
-35000.0
200
210
220
230
240
250
Wavelength (nm)
K3
K8
K7
K6


B
Fractional Folded
1
0.8
0.6
0.4
0.2
0
0
1
2
3
4
5
6
7
8
GdnHCl (M)
K3
K7
K8
K6


C
[θ]₂₂₂ (deg cm² dmol⁻¹)
0
-5000
-10000
-15000
-20000
-25000
-30000
10⁻⁷
10⁻⁶
10⁻⁵
0.0001
0.001
Concentration (M)
K3
K7
K8
K6

suggesting an increase in coiled-coil stability and association compared to the native kinesin sequence (K8 versus K3 in Fig. IV-6A).

To verify that the changes in helical content observed are in fact a result of changes in the stability of the respective coiled-coils, we determined the stability of K6, K7 and K8 by GdnHCl denaturation (Fig. IV-6, panel B and Table IV-1). The introduction of kinesin Y_{344}, E_{347}, and N_{351} into the model coiled-coil sequence caused a dramatic decrease in the stability (compare K6 and K7, GdnHCl midpoints. of >8 and 2.73 respectively). Correspondingly, the introduction of the three "ideal" model hydrophobic residues into the kinesin sequence dramatically increased the stability of the kinesin peptide K3 by 2.74 kcal/mol (GdnHCl midpoints of 1.17 M and 3.58 M for K3 and K8, respectively).

The changes in stability observed in Fig. IV-6 panel B are also reflected in the concentration dependence of the helical content, as measured at 222 nm (Fig. IV-6, panel C). The model coiled-coil sequence K6 exhibited practically no concentration dependence, which is consistent with its high degree of stability. Introduction of kinesin residues Y_{344}, E_{347}, and N_{351} into this sequence caused the coiled-coil now to dissociate upon dilution (compare K6 and K7, Fig. IV-6 panel C). Conversely, introduction of the "ideal" hydrophobes into the native kinesin sequence (K8) dramatically decreased its concentration dependence, which is consistent with its increased stability (compare K3 and K8, Fig. IV-6, panel B). Collectively, these data indicate that residues Y_{344}, E_{347}, and N_{351} destabilize the central region of the coiled-coil domain in kinesin. It is intriguing that both positions Y_{344} (Y or F) and E_{347} are very well conserved in the neck domains of the conventional kinesin, bimC/Eg5, and Kif3 (heterotrimeric) subfamilies of kinesin motors (Case, R. and Vale, R., manuscript in preparation).

Predictions of helical propensity and hydrophobicity in the kinesin neck

Finally, we wished to determine whether the observed location of the coiled-coil dimerization domain (res. 330-369) agrees with a predictive method developed in our own laboratory. The criteria of our method for predicting coiled-coil regions, which is very similar to that used by others, is based on the fact that coiled-coil regions are typically high in α–helical propensity (which exists over a minimum of 21 successive residues) as well as amphipathic, with hydrophobic residues occurring in a 3-4 repeating pattern. Analysis of the helical propensity of human kinesin sequence between residues 280 to 420 shows that there are three regions of high helical propensity that are well above our statistically determined cutoff value of 440 (unpublished results). These regions are indicated by the three connected boxes shown above the plot in Fig. IV-7 panel A. Of the three helical sections, only the region spanning residues 332-369 meets the minimum 21 successive residue cut off. This prediction agrees well with the experimentally observed stability of peptide K2 (residues 330-369). The residues adjacent to 369 dramatically drop in helical propensity, which agree with our experiments showing that residues 370-383 do not add any helical content to that of the region 330-369. Residues N-terminal to 332 also drop in helical propensity, which is again consistent with our results. It is interesting that the region from residue 325 to 340, which shows a large dip in helical propensity centered around residue 330, was predicted by Huang *et al.* (1994) to be high in β-sheet propensity based upon the secondary structure prediction program of Holly and Karplus (Holley *et al.*, 1989). The peptide K1 (316-355) which spans this region appears to adopt β-sheet secondary structure at pH 5.5.

Analysis of hydrophobicity occurring in a 3-4 repeating pattern (Figure IV-7, panel B) shows that there is only one dominant hydrophobic face in the region of high helical propensity. This face (B) contains the following hydrophobic residues (Fig. IV-1):

```
330                 347                  368
- C - - V- - - A- - - W - - - Y - - E - - - N - - L - - - I - - L - - - L - - W
```

The choice of these residues agrees with those previously predicted using the Lupas algorithm (Fig. IV-1, underlined residues) (Lupas *et al.*, 1991). The hydrophobicity of the heptad repeats appear to fluctuate dramatically. The greatest hydrophobicity occurs near the

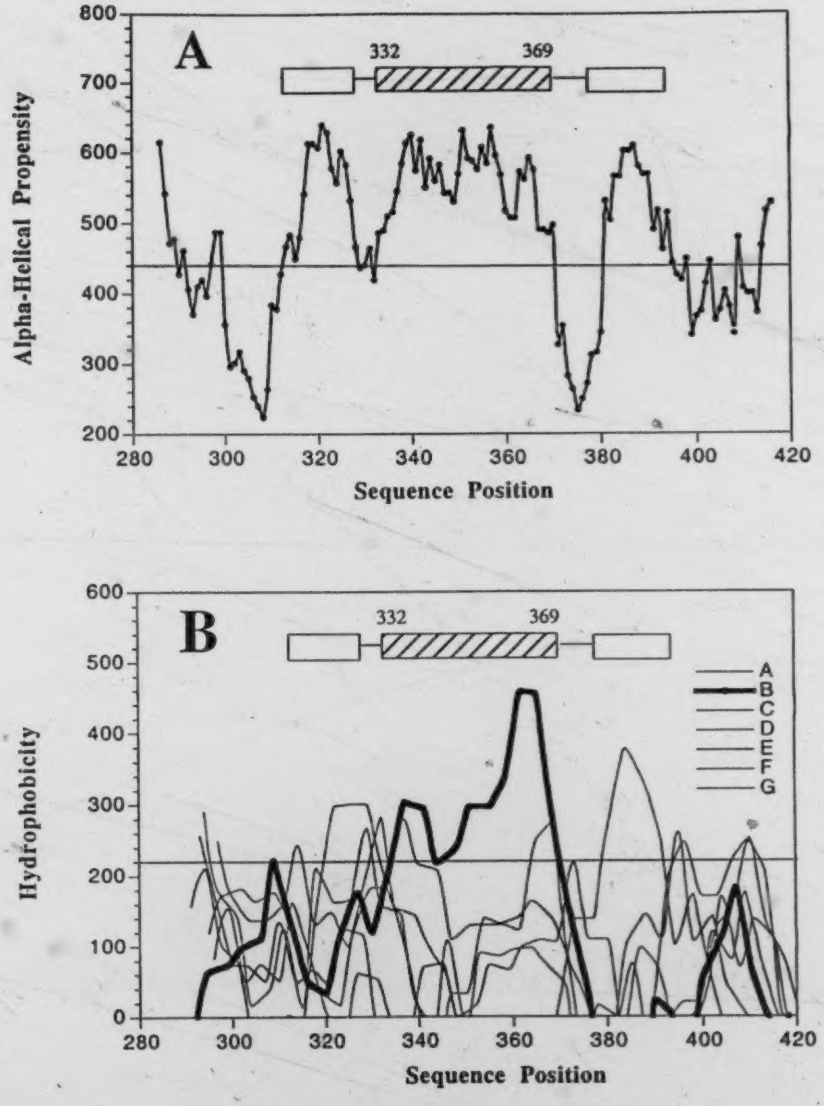


Figure IV-7. Helical propensity and hydrophobicity of the kinesin neck region. **Top,** Plot of the helical propensity from residues 280 to 420 of the human kinesin heavy chain . Each point represents the sum of 11 consecutive helical propensity scores. Regions which score greater than a pre-determined cut off value of 440 are indicated by the rectangular boxes. Regions which contains >21 consecutive residues (3 heptads) above the cut off value are indicated by hatching. **Bottom,** Plot of the hydrophobicity occurring in a 3-4 repeating pattern of the kinesin neck region from residues 280 to 420. Each data point represent the sum of the hydrophobicty scores for six consecutive "a" and "d" positions in a coiled-coil. All seven starting faces are shown. The face with the greatest hydrophobicity occurring in the region of high helical propensity is shown in bold. Face B corresponds to residues underlined in the native kinesin sequence in Figure IV-1.

C-terminus (hydrophobic residues of L, I, L, L, and W), an intermediate hydrophobicity is seen near the NH_2-terminus (residues C, V, A, and W), and the weakest hydrophobic heptad repeats are centered around Y, E and N (discussed previously). Although the results show fluctuating hydrophobicity within the proposed α-helical region, there is still significant hydrophobicity occuring in a 3-4 repeating pattern over a "large enough" range to predict that a coiled-coil could form. Thus, these results agree with those experimentally observed.

C. Discussion

Distinct subdomains of the kinesin neck

Using several synthetic peptides that overlap within the human kinesin heavy chain neck region (residues 316-383), we have been able to identify distinct subdomains with different secondary structure using CD spectroscopy. The central region from residue 330 to 369 of human kinesin heavy chain shows all of the characteristics of a stable two stranded α-helical coiled-coil. This region shows a significant α–helical spectrum with double minima at 222 and 208 nm, as well as a ratio of $[\theta]_{222}/[\theta]_{208}$ of 1.06, which is often indicative of such structures (Zhou *et al.*, 1992a,b; Zhu *et al.*, 1992). In addition, the location of this "coiled-coil" region correlates well with our own coiled-coil prediction method based on helical propensity and hydrophobicity. Stability, concentration dependence and gel filtration analyses have indicated that the 330-369 coiled-coil region forms a very stable and tightly associated dimer. Hence, we refer to the 330-369 region of the neck as the "dimerization domain." Furthermore, we have also been able to establish that the dimer is most likely oriented in a parallel fashion, since oxidation of cysteine 330 preserves the helical content and only enhances the stability. This would occur only if a disulfide bond was formed between two parallel-oriented chains. These results agree with previous reports that the kinesin heavy chains are arranged in a parallel and in-register orientation (de Cuevas *et al.*, 1992). Hence, this work conclusively shows that the kinesin neck contains a region that is capable of forming a two-stranded α–helical coiled-coil, as had been previously suggested by sequence predictions (Stewart *et al.*, 1993; Huang *et al.*, 1994).

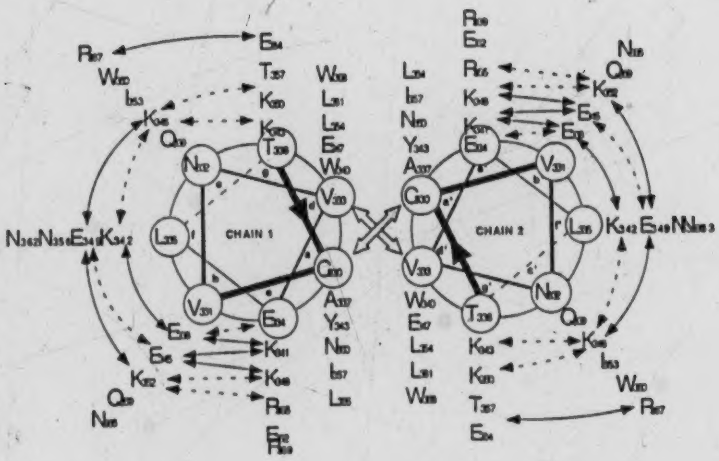
Further, the determined location of the dimerization domain also agrees with the results of Huang *et al.* (1994) and Correia *et al.* (1995), who showed that Drosophila kinesin heavy chains truncated to residue 392 and 401, respectively, form stable dimers (there residues are equivalent to human residues 384 and 393). In addition, similar findings with synthetic kinesin neck peptides to those described here have now also been obtained by Morii *et al.* (1997).

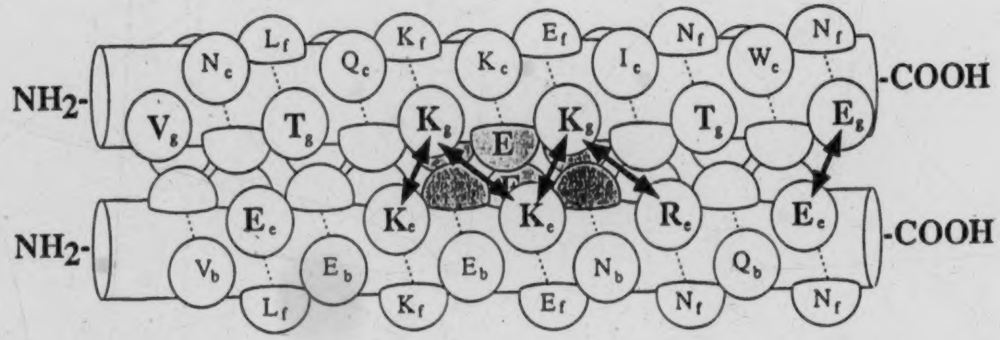
One of the questions raised from the previous work on truncated kinesin heavy chains was whether the putative coiled-coil in the neck region was sufficient on its own to account for dimerization, or whether other interactions (e.g. possibly between the head domains) were also involved. The estimated equilibrium dissociation constant for peptide 330-369 (4.3×10^{-9} M) is comparable to the value obtained for Drosophila K401 ($\sim4\times10^{-8}$ M) based upon equilibrium sedimentation studies (Correia *et al.*, 1995). Thus, our data indicates that the dimerization domain located within the neck is sufficiently stable on its own and thus could account for the dimerization observed in these recombinantly expressed motor domain constructs.

In contrast to the mostly helical dimerization domain, our work also indicates that the adjacent C-terminal (370-383) and N-terminal (320-332) segments show little helical propensity. The C-terminal residues 370-383 contribute little to dimer stability, and we suspect that these residues are part of a flexible hinge that is thought to exist between the neck and the stalk regions. The N-terminal segment of the neck, on the other hand, appears to have the capability of adopting a β-sheet secondary structure. This notion is based upon the finding that deletion of two C-terminal heptads from the coiled-coil dimer containing residues 316-369 (K5) causes the residual peptide (316-355) to display a CD spectrum characteristic of a β-sheet at pH 5.5 and to form a gel at neutral pH. We also observed that the addition of fourteen NH_2-terminal residues (316-329) onto the coiled-coil dimerization domain (330-369) caused a net loss of ~8 helical residues without decreasing the stability of the dimeric structure. This would suggest that a nonhelical intermolecular interactions (e.g. a small β-sheet secondary structure) can occur in the region between ~325-335. Our observation of β-sheet secondary structure also agrees with a secondary structure predictive algorithm of Holly and Karplus for the N-terminal portion of the neck by Huang *et al.* (1994). Since the N-terminal segment appears to be primarily non-helical and connects helix 6 in the crystal structure of the globular motor domain (Kull et al., 1996) to the helical dimerization domain, we refer to this segment of the neck as the "β linker region."

The "β linker region" of the neck is of considerable interest, since it appears to play an important role in motility and is highly conserved in many N-terminal motor proteins in the kinesin superfamily. Truncation at Drosophila kinesin residue 340 (human kinesin 332) eliminates the dimerization domain, and yet the motor still produces directional movement in a multiple motor microtubule gliding assay (Yang *et al.*, 1990; Stewart *et al.*, 1993; Berliner *et al.*, 1995; Vale *et al.*, 1996). Amino acid mutants in the "β linker region" of the neck also yield kinesin proteins that are severely defective in motility (unpublished observations). Whether a structural change occurs in the linker region during the force-generation cycle is unknown. However, the finding that this region can revert to a fully helical structure in the presence of 50% TFE indicates that this region has the intrinsic ability to adopt both β-strand and helical structures depending on the external environment. Thus, one could imagine that the "β linker region" could undergo a structural transition during the ATPase cycle, as will be discussed below.

Figure IV-8. End and side views of a two-stranded α-helical coiled-coil model for residues 330 to 369 of the kinesin neck region. **Panel A**, View from the NH_2-terminus, starting at residue 330. The direction of propagation of the peptide chains is into the page from N to C terminus with the chains parallel and in register. Residues in the first two helical turns are circled. Heptad positions are labeled "a" through "g", with the prime indicating corresponding positions on the opposing helix. Wide arrows depict the hydrophobic interaction that occur between residues in the "a" and "d" positions. Possible electrostatic interactions (attractions and repulsions) occurring on the outside of the helices are indicated as solid and broken arrows respectively. **Panel B**; Side view. Residues in positions b, c, e, f and g are shown. Residues in the e and g positions are in bold. Potential electrostatic interactions occurring across the hydrophobic face are indicated. The position for Tyr_{344}, Glu_{347} and Asn_{351} within the hydrophobic core are shaded.


CHAIN 1
CHAIN 2


NH2-
-COOH
NH2-
-COOH

Coiled-coil interactions within the dimerization domain are organized in a strong-weak-strong manner

GdnHCl denaturation studies of the kinesin peptides indicate that the coiled-coil dimerization domain is arranged in a strong-weak-strong pattern. Full stability of the dimerization domain is achieved only when all six of the spanning heptads (from residues 330-369) are present. Deletion of two heptads (14 residues) from the C- terminus (K1 peptide) results in a complete loss of all α-helical content, which is surprising considering that four of the six heptads still remain. Deletion of the two N-terminal heptads (K3 peptide), on the other hand, drastically decreases stability and ellipticity, although a dimeric structure can still be observed by gel filtration. These observations indicate that both the N- and C-terminal heptads are important for the stability of the structure, with the C-terminal heptads appearing to be the most important. These results agree with those of Corriea *et al.* (1995), who showed that truncation of the C-terminal 1.5 heptads of the proposed dimerization domain in *Drosophilia* kinesin produces a protein (K366) that fails to dimerize at concentrations up to 4 μM.

Interestingly, the central portion of the dimerization domain contains three residues, Y_{344}, E_{347} and N_{351}, that destabilize the coiled-coil structure. Introducing these amino acids into a model coiled-coil sequence has a significant destabilizing effect, and conversely, substituting these three residues in a kinesin neck peptide with "ideal" hydrophobic residues increases the stability of the coiled-coil interaction. Residues Y_{344} and E_{347} are highly conserved among several classes of N-terminal kinesin motors, suggesting that their role in destabilizing the central region of the coiled-coil may serve an important function. It is interesting to note that the E_{347} residues within the "hydrophobic" core are surrounded by opposite charged lysine and arginine residues in adjacent "e" and "g" positions (Fig. IV-8). Glover *et al.*, (1995) observed in a c-Fos c-Jun coiled coil crystal structure that a lysine residue located within a core position could potentially form hydrogen bonds and/or electrostatic interaction with adjacent residues in the "e" and "g" positions and speculated that this could stabilize the structure. Therefore the devastating effect of packing glutamic acid residues into the hydrophobic core in the kinesin neck could be mitigated to some extent by the formation of salt bridges with nearby residues. Such electrostatic interactions could possibly explain the higher than expected equilibrium association constant of the K3 peptide, even though it is very unstable in GdnHCl (which disrupts salt bonds).

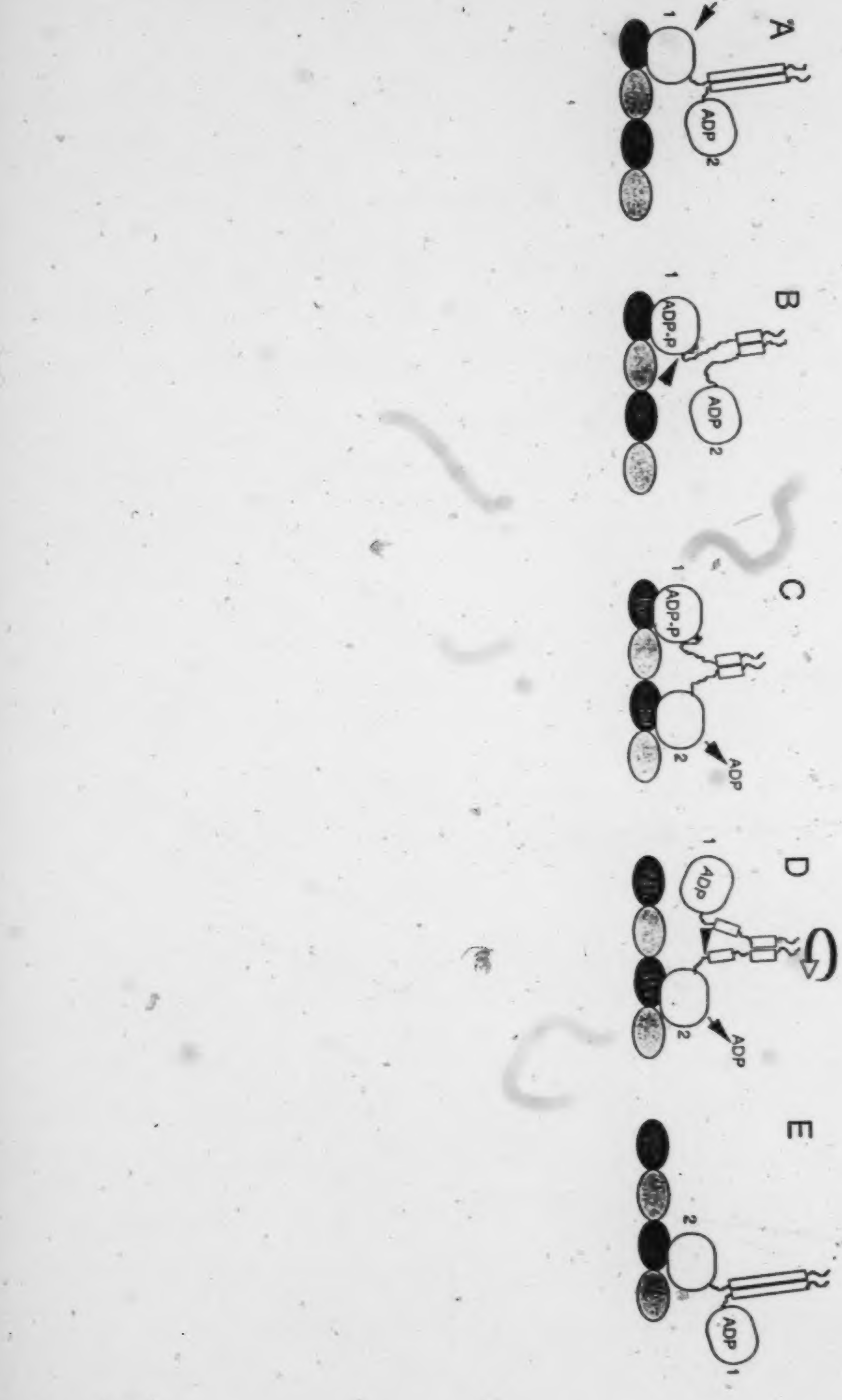
Another intriguing feature of the model representation of the coiled-coil dimerization domain in Fig. IV-8 is that the majority of electrostatic interactions across the core (e->g') are repulsive (indicated by the arrows in Fig. IV-8). Previous studies with model coiled-coil sequences have shown that attractive electrostatic interactions can be used to increase coiled-coil stability, control dimer orientation (parallel versus antiparallel), and govern homo versus heterodimerization (Zhu *et al.*, 1993; Kohn *et al*, 1995a,b; Monera *et al.*, 1993, 1994; Zhou 1992b,1994; for recent reviews, see also Adamson *et al.*, 1993; Alber, 1992; Baxevanis and Vinson, 1993; Hodges, 1992, 1996). The lack of significant attractive electrostatic interactions, taken together with our data showing an instability within the hydrophobic core, may indicate that the stability of the neck domain is optimized not only for its structure but also for its function as discussed below.

A model for kinesin motility involving structural transitions within the neck region

Studies on the kinesin dimer have indicated that the enzymatic cycles of the two kinesin motor domains may be coupled during processive movement. The strongest evidence for this idea comes from Hackney (1994) who showed that a kinesin dimer containing two bound ADP molecules releases ADP rapidly from one site and slowly from a second site after mixing with microtubules in a nucleotide-free buffer. Since microtubule interaction stimulates ADP release, Hackney suggested that after one head bound to the microtubule, the partner head had restricted access to a microtubule binding site. This idea is also consistent with recent cyro-EM images of microtubules decorated with dimeric kinesin, which show one kinesin head bound to the microtubule and the second head detached and oriented towards the plus-end of the microtubule (Arnal *et al.*, 1996; Hirose *et al.*, 1996). Hackney (1994) and Ma and Taylor (1997) have also shown that ATP binding and hydrolysis by the microtubule-bound kinesin head enables the partner head to bind to the microtubule and release its ADP. These results have led to the suggestion that the two kinesin heads in a dimer are predominantly in different states: one head strongly binds to the microtubule and weakly binds to ADP, while the second head weakly binds to the microtubule and strongly binds to ADP (Hackney (1994); Ma and Taylor (1996)). The two heads are suggested to alternate between these two states during processive movement.

Our results on the thermodynamic properties of kinesin neck peptides suggest a model that could provide an explanation for the results described above. The model shown in Fig. IV-9 begins with one head (without nucleotide) tightly bound to the microtubule, while the second head (containing ADP) is detached and directed towards the microtubule plus-end (Fig. IV-9, step A to B). In this state, the 6 heptad repeats in the neck domain are proposed to exist in a coiled-coil structure, which constrains the detached head from reaching an adjacent microtubule binding site. After the bound head binds and hydrolyzes ATP, a conformational change occurs in the globular motor domain, which is propagated to the nearby neck domain (Fig. IV-9, steps B to C). As a consequence, the "β linker region" of the neck domain changes its structure or interacts in a new manner with the core motor domain, and thereby acts like a contracting spring. This conformational change would account for the observation that monomeric kinesin containing the "β linker region," although not processive, can elicit force and directional movement. In the kinesin dimer, the conformational change in the "β linker region" results in a loss of α-helical residues at the NH_2-terminus of the coiled-coil dimerization domain. This, in turn, could result in a cooperative unzippering of the majority of the coiled-coil up to the last two stable heptad repeats (since the middle two heptads are inherently unstable). Unwinding of the coiled-coil would create a flexible linker between the heads, which would allow the detached head

Figure IV-9. Schematic representation of the possible conformational changes occurring in the neck region during kinesin movement. Globular spheres depict the motor domain heads, rectangular boxes indicate the kinesin neck region which forms a two stranded α-helical coiled-coil, and the shaded ovals at the base indicate the α and β tubulin subunits. In step A, one motor domain head of kinesin is tightly bound to the microtuble filament while the second head (containing ADP) is detached and directed toward the "plus" end of the microtubule, as suggested by cryo-electron microscopy studies. The two heads may be related by ~180° symmetry (Howard, 1996). After hydrolysis of ATP by the bound motor domain, a conformational change is transmitted from the active site to the β linker region (arrow, step B) which causes the first four heptads of the neck region to "unzip". (ADP-Pi is shown in the active site, although this change could equally well occur after phosphate release). The unzipping of several heptads in the coiled-coil allows the unbound head to now reach the next available tubulin subunit (Steps B to C). Release of ADP from the newly bound motor domain head to create a tight microtubule binding state coupled with the progression of the other head to a weak microtubule binding state reverses the structural changes in the neck region and enables the coiled-coil dimerization domain to "re-zip" (Steps C to D). Coupled to re-zipping, is the rotation of the detached head by 180° (circular arrow, Howard, 1996). The kinesin enzyme then returns to its initial state (Step E). For a more detailed description of each step occuring in the model see the "Discussion section".


A
ATP
1
ADP
2
B
1
ADP-P
ADP
2
C
1
ADP-P
2
ADP
D
1
ADP
2
ADP
E
2
ADP
1

to reach an adjacent microtubule binding site (step B to C). Directional movement to a forward binding site could be favored by the initial positioning of the detached head closer to the plus-end of the microtubule (Arnal *et al.*, 1996; Hirose *et al.*, 1996). Upon docking to the microtubule, the forward head would then release its ADP and enter a strong microtubule-binding state, while the rearward head would have progressed in the cycle to a weak-binding ADP state (Fig. IV-9, steps C to D). These events would also reverse the structural change in the linker region of the neck, which would permit the coiled-coil in the dimerization domain to "rezip". In the process, the rearward head would rotate by ~180° (Howard, 1996), returning the enzyme to its initial state in the cycle and thereby completing one mechanical step. The above model makes several predictions concerning the roles of different regions within the neck domain. Most notably, enhancing the stability of the middle heptad repeats in the dimerization domain should impair processivity and alternating head ATP catalysis as a result of increasing the energetic requirement for unravelling the coiled-coil as a prelude to separating the heads. Experiments are currently underway to examine this question.

D. References

Adamson, J. G., Zhou, N. E., and Hodges, R. S. (1993) *Curr. Opin. Biotechnol.*, **4**, 428-437.

Alber, T. (1992) *Curr. Opin. in Genet. & Dev.*, **2**, 205-210.

Arnal, I., Metoz, F., DeBonis, S. and Wade, R. H. (1996) *Curr. Biol.* **6**: 1265-1270.

Baxevanis, A. D., and Vinson, C. R. (1993) *Curr. Opin. in Gen. and Dev.*, **3**, 278-285.

Berliner, E., Young, E. C., Anderson, K., Mahtani, H. and Gelles, J. (1995) *Nature*, **373**: 718- 721.

Bloom, G. S., Wagner, M. C., Pfister, K. K., and Brady, S. T. (1988) *Biochemistry*, **27**, 3409- 3416

Chen, Y.H., Yang, J.T., and Martinez,H.M. (1972) *Biochemistry*, **11**, 4120-4131.

Chen, Y.H., Yang, J.T., and Chau, K.H. (1974) *Biochemistry*, **13**, 3350-3359.

Correia, J.J., Gilbert, S.P, Moyer, M.L., and Johnson, K.A. (1995) *Biochemistry*, **34**, 4898-4907.

de Cuevas, M., Tao, T., and Goldstein, L. S. B. (1992) *J. Cell Biol.* **116**, 957-965.

De Francesco, R., Pastore, A., Vecchio, G., and Cortese, R. (1991) *Biochemistry*, **30**, 143-147.

Engel, M., Williams, R. W., and Erickson, B. W. (1991) *Biochemistry*, **30**, 3161-3169.

Farrow, N. A., Muhandiram, R., Singer, A.U., Pascal,S.M., Kay, C.M., Gish,G., Shoelsen, S.E., Pawson, T., Forman-Kay, J.D., and Kay, L.E.(1994) *Biochemistry* **33**, 5984-6003.

Gans, P.L., Lyu, P.C., Manning, M.C., Woody, R.W., and Kallenbach, N.R. (1991) *Biopolymers*, **31**, 1605-1614.

Glover, J.M.N., and Harrison, S.C. (1995) *Nature*, **373**, 257-261.

Graddis, T. J., Myszka, D. G., and Chaiken, I. M. (1993) *Biochemistry*, **32**, 12664-12671.

Hackney, D. D. (1995). Nature **377**, 448-450.

Hackney, D. D. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**:6865-6869.

Hirokawa, N., Pfister, K. K., Yorifuji, H., Wagner, M. C., Brady, S. T., and Bloom, G. S. (1989) *Cell*, **56**, 867-878.

Hirose,K., Lockhart, A., Cross, R.A., Amos, L.A. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9539-9544.

Hodges, R. S. (1992) *Curr. Biol.*, **2**, 122-124.

Hodges, R.S. (1996) *Biochem. Cell Biol.*, **74**, 133-154.

Holley,L.H. and Karplus, M. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 152-156.

Howard, J. 1996. *Annu. Rev. Physiol.* **58**, 703-729.

Huang, T.G., Suhan, J., Hackney, D.D. (1994) *J. Biol. Chem.* **269**, 16502-16507.

Kohn, W. D., Kay, C. M., and Hodges, R. S. (1995a) *Prot. Sci.*, **4**, 237-250.

Kohn, W. D., Monera, O. D., Kay, C. M., and Hodges, R. S. (1995b) *J. Biol. Chem.*, **270**, 25495-25506.

Kull, F.J., Sablin, E.P., Lau, R., Fletterick, R.J., and Vale, R.D. (1996) *Nature*, **380**, 550-555.

Kuznetsov, S. A., Vaisberg, E. A., Shanina, N. A., Magretova, N. N., Chernyak,V.Y., and Gelfand, V. I. (1988) *EMBO J.* **7**, 353-356.

Lau, S. Y. M., Taneja, A. K., and Hodges, R. S. (1984b) *J. Chromatogr.*, **317**, 129-140.

Lupas, A., Dyke, M.V., and Stock, J. (1991) *Science*, **252**, 1162-1164.

Ma, Y. -Z. and Taylor, E. W. (1997) *J. Biol. Chem.* (in press).

Monera, O. D., Zhou, N. E., Kay, C. M., and Hodges, R. S. (1993) *J. Biol. Chem.*, **268**, 19218-19227.

Monera, O. D., Kay, C. M., and Hodges, R. S. (1994) *Biochemistry*, **33**, 3862-3871.

Monera, O.D., Sereda, T.J., Zhou, N.E., Kay, C.M. and Hodges, R.S. (1995) *J. Pept. Sci.*, **1**, 319-392.

Morii, H., Takenawa, T., Arisaka, F., and Shimizu, T. (1997) *Biochemistry* (in press).

O'Shea, E. K., Klemm, J. D., Kim, P. S., and Alber, T. (1991) *Science*, **254**, 539-544.

Pace, C.N., (1986) *Methods Enzymol.* **131**, 266-279.

Sablin, E. P., Kull, F.J., Cooke, R., Vale, R.D. and Fletterick, R.J. (1996) *Nature*, **380**, 555-559.

Scholey, J. M., Heuser, J., Yang, J. T., and Goldstein, L. S. B. (1989) *Nature*, **338**, 355-357

Schuermann, M., Hunter, J. B., Hennig, G., and Muller, R. (1991) *Nucleic Acids Res.*, **19**, 739-746.

Sereda, T. J., Mant, C. T., Quinn, A. M., and Hodges, R. S. (1993) *J. Chromatogr.*, **646**, 17-30.

Shortel, D. (1989) *J.Biol. Chem.* **264**, 5315-5318.

Stewart, R.J., Thaler, J.P., and Glodstein, L.S.B. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 5209-5213

Sonnichsen, F.D., Van Eyk, J.E., Hodges, R.S. and Sykes, B.D. (1992) *Biochemistry*, **31**, 8790-8798.

Su, J. Y., Hodges, R. S., and Kay, C. M. (1994) *Biochemistry*, **33**, 15501-15510.

Vale, R. D. 1993. Motor Proteins. *in* Guidebook to the Cytoskeletal and Motor Proteins, Kreis, T., Vale, R., eds. (Oxford Univ. Press). pg. 175-183.

Vale, R.D., Funatsu, T., Pierce, D. W., Romberg, L., Harada, Y., and Yanagida, T. (1996) Nature **380**: 451-453.

Vinson, C. R., Hai, T., and Boyd, S. M. (1993) *Genes and Dev.*, **7**, 1047-1058.

Yang, J. T., Laymon, R. A., and Goldstein, L. S. B. (1989) *Cell*, **56**, 879-889

Yang, J. T., Saxton, W. M., Stewart, R. J., Raff, E. C.and Goldstein, L. S. (1990). *Science*, **249**, 42-47.

Young, E.C., Berliner, E.,Mahtani, H.K., Perez-Ramirez, B., and Gelles, J. (1995) *J. Biol. Chem.* **270**, 3926-3931.

Zhou, N. E., Kay, C. M., and Hodges, R. S. (1992c) *J. Biol. Chem.*, **267**, 2664-2670.

Zhou, N. E., Zhu, B. Y., Kay, C. M., and Hodges, R. S. (1992b) *Biopolymers*, **32**, 419-426.

Zhou, N. E., Kay, C. M., and Hodges, R. S. (1992a) *Biochemistry*, **31**, 5739-5746.

Zhou, N.E., Kay, C.M., and Hodges, R.S. (1993) *Biochemistry*, **32**, 3178-3187.

Zhou, N. E., Kay, C. M., and Hodges, R. S. (1994) *J. Mol. Biol.*, **237**, 500-512.

Zhu, B. Y., Zhou, N. E., Kay, C. M., and Hodges, R. S. (1993) *Prot. Sci.*, **2**, 383-394.

Zhu, B. Y., Zhou, N. E., Semchuk, P. D., Kay, C. M., and Hodges, R. S. (1992) *Int. J. Pept. Prot. Res.*, **40**, 171-179.

Chapter V

Investigation of the Structural Features within the Human Kinesin Neck Region.

A version of this chapter is presently being revised for publication

A. Introduction

Kinesin is a dimeric motor protein that can move for several micrometers along a microtubule without dissociating. The two kinesin motor domains are thought to move processively by operating in a "hand-over-hand" manner, although the exact mechanism underlying such coordinated cooperatively is still unknown (Block, 1998; Mandelkow and Johnson, 1998). Recently, an amino acid region comprised of *ca.* 50 residues adjacent to the globular motor domain head (termed the neck) has been shown to be sufficient for conferring dimerization and processive movement. Based upon its amino acid sequence, the neck was proposed to dimerize through a coiled-coil interaction. Recent structural studies using synthetic peptides and X-ray crystallography have indeed shown this to be the case (Kozielski *et al.*, 1997; Morri *et al.*, 1997; Tripet *et al.*, 1997). The neck region consists of an N-terminal region, comprised of *ca.* 10 amino acid residues, which forms a intramolecular beta-sheet secondary structure (termed the β-linker) followed by a 32 amino acid residue α-helix which forms a parallel two-stranded α-helical coiled-coil (termed the neck coiled-coil).

One of the most interesting features highlighted by the recent structural analysis of the kinesin neck region has been the presence of several highly conserved but "destabilizing" residues existing within the first two heptads of the coiled-coil region. Hence, this has raised speculation whether the neck region coiled-coil not only enables head-to-head communication by creating a stiff connection linkage between the two motor domains in one nucleotide binding state, but may also 'unfold' or 'unwind' in another nucleotide binding state to allow movement to the next tubulin binding site. At present, several models describing such a mechanism have recently been reported (Hackney *et al.*, 1994, 1995; Tripet *et al.*, 1997; Mandelkov *et al.*, 1998).

In an effort to determine the validity of this hypothesis, two groups have recently investigated the importance of the destabilizing residues within the neck region coiled-coil or those residues which specifically interface with the coiled-coil structure and the motor domain. For example, Romberg *et al.*, (1998) have recently substituted the 'destabilizing' hydrophobic core residues Tyr 344, Glu 347 and Asn 351 with Ile, Leu and Ile, respectively, or completely replaced the coiled-coil region C-terminal from residues 343 to 370 with a more stable coiled-coil structure [(EIEALKA)4; STABLECOIL] to observe the effect upon single motor motility assays. Interestingly, their results showed that, in contrast to the expected abolition of processive activity by preventing the separation of the catalytic motor domain heads, the activities were in fact reduced only marginally (2 fold). Even more interesting, a motor domain construct which contained a repeated sequence of the first heptad of the neck region coiled-coil (DUP; residues 336 to 342; TAEQWKK) showed enhanced motor activity and levels of processivity. As such, the authors concluded that separation of the adjacent neck coiled-coil region is not an absolute requirement for coordinated kinesin motor processivity *per se*, however, structural features within the neck region coiled-coil are important for modulating or regulating this activity. Subsequent results by Hackney (Jiang *et al.*, 1999) who extended a spacer between the coiled-coil and the motor domain and still observed motor processivity, also appear to be in agreement with such a proposal. Yet, in spite of this apparent congruency of results, exactly which structural region is "unfolding" and generating the required 8 nm distance in order to attach to the next microtubule binding site in the "native state" is still a controversial topic (Mandelkov and Johnson, 1998).

Previously, we (Tripet, *et al.*, 1997; Chapter IV) demonstrated the utility of using several synthetic peptide analogs corresponding to the human kinesin neck region in order to establish clearly the presence of the coiled-coil dimerization domain, its relative location and its relative stability. From these studies we predicted the neck region coiled-coil spanned ~330-369 (human kinesin sequence). More recently, the X-ray crystal structure of monomeric and dimeric rat brain kinesin has been solved and clearly shows the neck coiled-coil structure to begin at residue 339 (337 of the human sequence) and proceed to approximately residue 370 (368 of the human sequence) (Kozielski, *et al.*, 1997; Sack, *et al.*, 1997). In comparison with our earlier studies, the starting residue of the coiled-coil is 6 residues closer to the C-terminus of our mapped region, and thus we have analyzed the coiled-coil region in the context of the small beta-sheet stretch (β10; residues 329 to 333; CVNV) (Sack, *et al.*, 1997) which we have speculated has an important effect on the overall stability of the coiled-coil region. Additionally, the last two heptads of the coiled-coil region have been noted to show sequences similar to that of a 'trigger sequence' important for the induction and/or nucleation of coiled-coils (Kammerer, *et al.*, 1998; Steinmetz, *et al.*, 1998). In light of a possible unfolding mechanism, it has thus been speculated that the last two C-terminal heptads may be playing a critical role in triggering the transient re-folding of the coiled-coil should it be separating during movement. Although our previous studies demonstrated that the absence of the last two heptads resulted in a peptide which displayed the complete absence of helical structure, thus supporting such a proposal (Tripet, *et al.*, 1997), we did not synthesize a specific analog in which only the last heptad was deleted in order to support or dispute such a proposal.

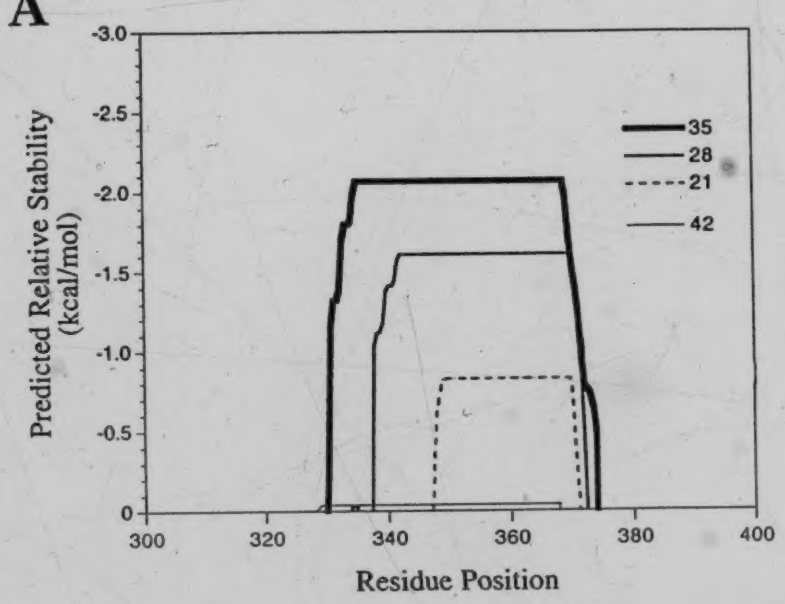
Thus, in order to investigate these issues further, as well as clearly establish the presence or absence of any other unique structural features within the kinesin neck region, we have extended our original peptide mapping study (Chapter IV) to include several new peptide variants of the human kinesin neck region. Following extensive characterization of these analogs (CD spectroscopy and GdnHCl denaturation studies), we have been able to develop a model relating the function of the various structural features identified with a mechanism for kinesin processivity.

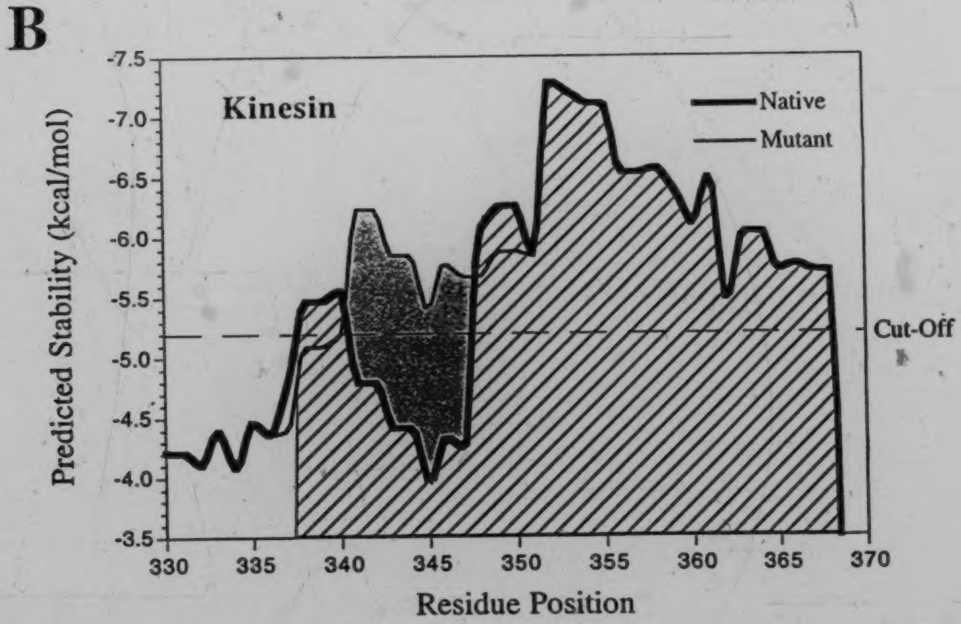
B. Results

Coiled-coil prediction and analysis of the human kinesin neck region.

In a related communication (Tripet, *et al.*, 1999), we have recently descibed the development of a new method for predicting the location and stability of putative coiled-coil dimerization domains within native protein sequences. This method, which differs from those previously described (Berger, *et al.*, 1995; Fischetti, *et al.*, 1993; Lupas, *et al.*, 1991), is based on the premise that the location of a putative coiled-coil domain can be accurately predicted by identifying regions within the native protein sequence that are sufficiently stable to form such structures. Stability in this regard results from (1) the intrinsic α-helical propensity (ability to adopt an α-helix) of the local region and (2) the ability to associate two or more α-helices in a stable manner through hydrophobic packing of the **a** and **d** residues of the heptad repeat (denoted **abcdefg**). Experimental data sets for each of these parameters (see methods) have recently been generated by this laboratory and have been used to create a scoring algorithm which initially assigns, then iteratively sums, the sequence regions for their "coiled-coil" stability using various sized window-widths and cut-off values. Results from the method to date have indicated that this method is as accurate or better than existing coiled-coil prediction methods with three additional benefits: (1) the method calculates the propensity of a sequence region to form a coiled-coil structure based entirely on its intrinsic physiochemical stability properties, thus making the predicted propensity value independent of any previously observed positional preference or statistical bias; (2) it can generate a "moving" stability profile for the predicted region(s) which can be used to define more accurately the exact termini of such regions; and (3) variations within the moving stability profile are able to illuminate important and/or interesting structural features within such region(s) that may be important for the structure and/or functioning of the domain.

Figure V-1. *Panel A*; Prediction of α-helical coiled-coil formation within the human kinesin neck region (residues 300-400). Prediction was carried out as that described by the method of Tripet *et. al.* 1999 using stability data. The horizontal axis indicates residue position, and the vertical axis represents predicted relative stability. Only stability values which exceed the pre-defined cut-off value for the respective window sizes (21, 28, 35, and 42 residues) are plotted. The 35-residue region representing residues 332-372 is predicted with the greatest stability for coiled-coil formation. *Panel B*; The predicted 7-residue window width stability profile of the human kinesin neck region coiled-coil (defined above, panel A) used to define the exact start and end points of the coiled-coil as well as demonstrate the variation in the stability along the coiled-coil. The stability profile of the native human kinesin sequence is denoted in bold outline. The region predicted to exist as a coiled-coil using the 7-residue window width cut-off value (-5.2) is hatched. The increase in stability as a result of the substitution of K343, E347, and K350 in the native sequence with E343, L347 and E350 (kinesin peptide K2b, see Figure 2) is shown in solid grey. The mutant sequence stability profile is shown in the unbolded line.


A
Predicted Relative Stability (kcal/mol)
-3.0
-2.5
-2.0
-1.5
-1.0
-0.5
0
300
320
340
360
380
400
Residue Position
35
28
21
42


B
Kinesin
Native
Mutant
Cut-Off
Predicted Stability (kcal/mol)
-7.5
-7.0
-6.5
-6.0
-5.5
-5.0
-4.5
-4.0
-3.5
330
335
340
345
350
355
360
365
370
Residue Position

As described above, the neck region of kinesin is envisioned to possess some unique structural features important for motor processivity. In an attempt to highlight such features, we have now employed our new coiled-coil predictive program to re-analyze the kinesin neck region coiled-coil. Figure V-1 shows the generated output for the scoring of the kinesin neck sequence from residues 300 to 400. Residues 330 to 372 exceed the predefined cut-off values and are thus predicted to contain sufficient intrinsic stability to form a coiled-coil structure (Fig. V-1A). Analysis of this region using the four window sizes 21, 28, 35 and 42 indicates the region is predicted to have the greatest stability using a window width of 28 and 35 residues; thus, the relative size of the coiled-coil region is predicted to be between 28 and 35 residues.

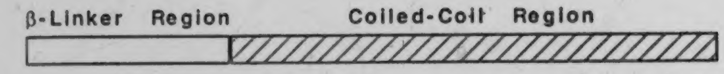
To more accurately define the N- and C-termini of the predicted coil-coil region, a 7-residue window width score was calculated and used to generate a "moving stability profile" for the predicted region (Fig. V-1B). In this plot the position at which the predicted stability plot initially exceeds the pre-defined cut-off value of 5.2 kcal/mol occurs at residue 337, and terminates the predicted region at residue 369. Thus, the coiled-coil is predicted to be 33 residues in length and span residues 337-369. Output from the 7-residue window width further illustrates that the stability of this region is organized in the following 'stability' manner: the first heptad (residues 337-343) is predicted to be of relatively low stability, i.e., it only just exceeds the pre-defined cut-off value, suggesting that the first heptad does not make a significant contribution based on its helical propensity and hydrophobicity; the second heptad (residues 341-347) is of extremely low stability (due in part to the destabilizing core residues of Tyr 344 and Glu 347, Tripet, *et al.*, 1997); and finally the last 3-heptads (residues 348-369) form a continuous region of high predicted stability.

Taken together, the prediction region appears to be in excellent agreement with both our earlier prediction and analysis of this region (Tripet, *et al.*, 1997) as well as with the recently solved X-ray crystallographic structure of rat brain dimeric kinesin (Kozielski, *et* *al*., 1997). Interestingly, however, the only apparent unique structural feature highlighted by the present program is the unusually low stability of the two N-terminal heptads.

Contribution of the N- and C-terminal heptads to stability and dimerization of the human kinesin coiled-coil.

Previously Tripet, *et al*., (1997) showed that deletion of the last two C-terminal heptads of the kinesin neck region coiled-coil resulted in a peptide which completely lacked α-helical secondary structure (Fig. V-3A; kinesin peptide K1). These results suggested that the last two heptads clearly play an important role in the induction and/or stabilization of the coiled-coil domain. In agreement with these findings, the latter two heptads have been noted for their sequence homology to a 'trigger sequence' important for inducing/nucleating coiled-coil formation (Krammer et al., 1998). Thus, to establish clearly the location of the C-terminus of the neck coiled-coil region in solution as well as demonstrate the importance of the last two C-terminal heptads for helix induction, we synthesized and characterized the kinesin peptide variant K1b (sequence shown in Fig. V-2). This peptide corresponds to residues 316-363 which, in comparison to our previous kinesin peptides K1 and K4 (Fig. V-2), is one heptad longer than that of K1 (previously observed to exhibit beta-sheet secondary structure, Fig. V-3A) and one heptad shorter than K4 (previously observed to exhibit α-helical conformation, Fig. V-3A).

Far UV circular dichroism (CD) spectroscopy analysis of the K1b peptide (Fig. V-3A) shows a characteristic α-helical spectrum with double minima at 208 and 222 nm . The degree of mean residue molar ellipticity (-11,068°) and ratio of $[\theta]_{222}/[\theta]_{208}$ (0.74) indicates that although this peptide adopts an α-helical conformation, it is unable, in contrast to peptide K4 (Tripet *et al.*, 1997), to form a fully folded dimeric coiled-coil structure as previously observed for the K4 peptide. In agreement with these observations, stability analysis of the K1b peptide, as assessed by GdnHCl denaturation (Fig. V-3B), demonstrated that deletion of the last 6 residues of the C-terminal heptad



316 330 335 340 347 355 369 383

fgabcdefg abcdefg abcdefg abcdefg abcdefg

Human Kinesin			LLFGQRAKTIKNTVCVNVELTAEQWKKKYEKEKEKNKILRNTIQWLENELNRWRNGETVPIDEQFDKE
*K1	(316-355)		Ac-LLFGQRAKTIKNTVCVNVELTAEQWKKKYEKEKEKNKILR-NH_2
K1b	(316-363)		Ac-LLFGQRAKTIKNTVCVNVELTAEQWKKKYEKEKEKNKILRNTIQWLEN-NH_2
*K4	(316-369)		Ac-LLFGQRAKTIKNTVCVNVELTAEQWKKKYEKEKEKNKILRNTIQWLENELNRWR-NH_2
*K5	(330-383)		Ac-CVNVELTAEQWKKKYEKEKEKNKILRNTIQWLENELNRWRNGETVPIDEQFDKE-NH_2
K3b	(337-383)		Ac-AEQWKKKYEKEKEKNKILRNTIQWLENELNRWRNGETVPIDEQFDKE-NH_2
*K3	(344-383)		Ac-YEKEKEKNKILRNTIQWLENELNRWRNGETVPIDEQFDKE-NH_2
*K2	(330-369)		Ac-CVNVELTAEQWKKKYEKEKEKNKILRNTIQWLENELNRWR-NH_2
K2b	(330-369)	(K343E, E347L, K350E)	Ac-CVNVELTAEQWKK**E**YEK**L**KE**E**NKILRNTIQWLENELNRWR-NH_2
K2c	(337-369)	(K343E, E347L, K350E)	Ac-AEQWKK**E**YEK**L**KE**E**NKILRNTIQWLENELNRWR-NH_2
K2e	(330-369)	(L335G, K343E, E347L, K350E)	Ac-CVNVE**G**TAEQWKK**E**YEK**L**KE**E**NKILRNTIQWLENELNRWR-NH_2
K2f	(330-369)	(L335G, T336G, K343E, E347L, K350E)	Ac-CVNVE**GG**AEQWKK**E**YEK**L**KE**E**NKILRNTIQWLENELNRWR-NH_2
K2d	(337-369)	(W340A)	Ac-AEQ**A**KKKYEKEKEKNKILRNTIQWLENELNRWR-NH_2
K2g	(330-369)	(W340A)	Ac-CVNVELTAEQ**A**KKKYEKEKEKNKILRNTIQWLENELNRWR-NH_2

Figure V-2. Amino acid sequences of the synthetic peptides used in this study. Native human kinesin sequence is shown at the top. Peptide names are indicated on the left. Numbers in parentheses indicate the amino acid sequence region which the peptide spans as well as the location of amino acid substitutions that differ from the native sequence. Ac- denotes N$^{\alpha}$-acetyl, and -amide denotes C$^{\alpha}$-amide. The position of all kinesin peptides are shown relative to the native sequence at top. Residues identified or proposed to be involved in forming the 3-4 hydrophobic repeat of the coiled-coil domains are underlined. The corresponding secondary structure regions (β-linker and coiled-coil) observed by X-ray crystallography are shown at the top of the sequence in a bar diagram. * denotes sequences that were reported previously (Tripet *et. al.* 1997) and are shown for direct comparison with the new peptides prepared in this study.

Figure V-3. *Panel A*; CD spectra of kinesin peptides K1, K1b and K4; Spectra for K1b and K4 peptides were recorded at 20°C in a 0.1 M KCl, 0.05 M PO_4, 0.002 M DTT, pH 7 buffer. The K1 peptide was recorded in the same buffer but at a pH of 5.5. Peptide concentrations were 100 µM. *Panel B;* GdnHCl denaturation profiles of kinesin peptides K1b and K4. Denaturation profiles were recrded at 20°C in 0.1 M KCl, 0.05 M PO_4, 0.002 M DTT, pH 7 buffer with various GdnHCl denaturant concentration. The fraction folded (f*f*) of each peptide was calculated as $f_f = ([\theta]-[\theta]_u)/([\theta]_n-[\theta]_u)$, where $[\theta]$ is the observed mean residue ellipticity at 222 nm at any particular denaturant concentration and $[\theta]_n$ and $[\theta]_u$ are the mean residue ellipticities at 222 nm of the native "folded" and "unfolded" states, respectively. Peptide concentrations were 60 µM. *Panel C*; CD spectra of kinesin peptides K3, K3b and K5; spectra were recorded at 20°C in a 0.1 M KCl, 0.05 M PO_4, 0.002 M DTT buffer, pH 7. Peptide concentrations were 100 µM. *Panel D*, GdnHCl denaturation profiles of kinesin peptides K3, K3b and K5. GdnHCl denaturation profiles and calculation of fraction folded (f*f*) were carried out as described above. Peptide concentrations were 60 µM. Note; CD spectra and denaturation profiles for kinesin peptides K1, K3, K4, and K5 have been reported previously (Tripet *et. al.* 1997) but are shown for direct comparison to the two new kinesin peptides K1b and K3b.

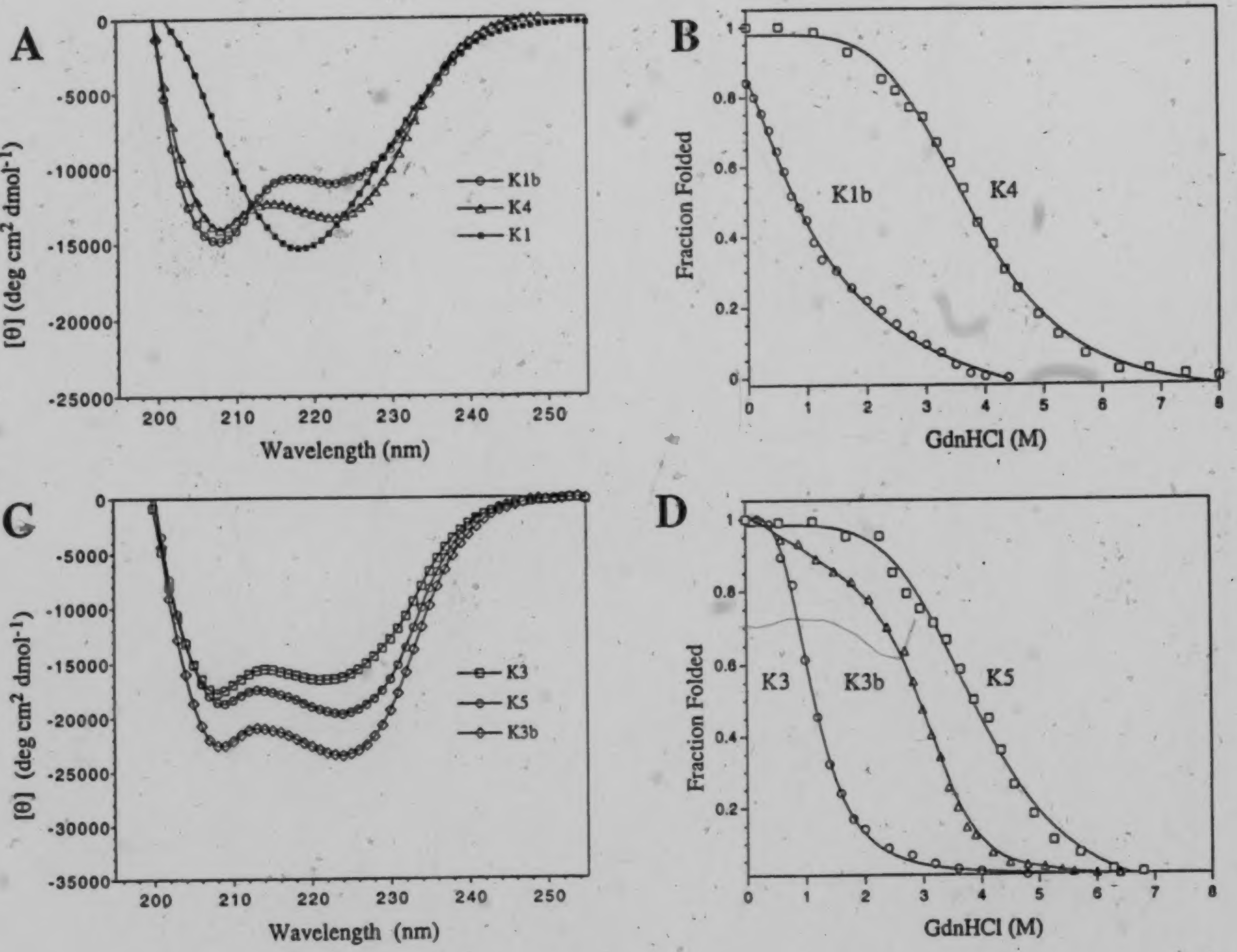

A
[θ] (deg cm² dmol⁻¹)
0
-5000
-10000
-15000
-20000
-25000
200
210
220
230
240
250
Wavelength (nm)
K1b
K4
K1
B
Fraction Folded
1
0.8
0.6
0.4
0.2
0
0
1
2
3
4
5
6
7
8
GdnHCl (M)
K1b
K4
C
[θ] (deg cm² dmol⁻¹)
0
-5000
-10000
-15000
-20000
-25000
-30000
-35000
200
210
220
230
240
250
Wavelength (nm)
K3
K5
K3b
D
Fraction Folded
1
0.8
0.6
0.4
0.2
0
1
2
3
4
5
6
7
8
GdnHCl (M)
K3
K3b
K5

Table V-1. Ellipticities and stabilities of the synthetic peptides

Peptide[a]	# of residues/ peptide chain	$[\theta]_{222}$ [b]		Helical Content[c]		$[\theta]_{222/208}$ [d]	$[\mathrm{GdnHCl}]_{1/2}$[e]	m[f]	$\Delta\Delta G_u$[g]
		Benign	50%TFE	(%)	#Res.				
		degrees · cm² · dmol⁻¹							*kcal · mol⁻¹*
K1b	48	-11,068	-25,858	31	15	0.74	0.82	2.4	-4.11
K4*	54	-13,300	-18,440	37	20	0.94	3.73	1.1	1.05
K2*	40	-24,600	-26,600	70	28	1.06	3.61	1.2	0.92
K2b	40	-29,255	-31,660	83	33	1.01	4.75	1.4	2.48
K2c	33	-29,255	-29,255	85	28	1.02	3.60	1.4	0.96
K3*	40	-16,500	-19,190	46	19	0.94	1.17	2.3	-3.33
K3b	47	-23,598	-23,967	66	31	1.03	2.93	1.5	0.00
K5*	54	-19,500	-21,200	54	29	1.05	3.93	1.1	1.32
K2e	40	-26,814	-28,444	76	30	0.96	3.71	1.1	1.01
K2f	40	-25,970	-32,238	73	29	0.95	3.35	1.2	0.57
K2d	33	-21,791	-35,223	63	21	0.90	0.50	1.6	-3.77
K2g	40	-23,111	-29,333	65	26	0.96	1.03	1.5	-2.85

Table V-1 Legend

a Amino acid sequences for each peptide are shown in Fig. 2.

b The mean residue molar ellipticities at 222 nm were measured at 20°C in benign buffer (0.1 M KCl, 0.05 M PO_4, pH 7). For samples containing TFE, the above buffer was diluted 1:1 (v/v) with TFE. Kinesin peptide concentrations were 100 μM. Kinesin-like peptide concentrations varied from 100-250 μM; see Figure legends for exact concentrations.

c The helical content (%) was calculated from the ratio of the observed $[\theta]_{222}$ value divided by the predicted molar ellipticity x 100. The predicted molar ellipticity was calculated from the equation $[\theta]_{222} = [\theta]^{\alpha} (1-k/N)$ for the chain length dependence of an α-helix (Gans *et. al* 1991), where $[\theta]^{\alpha} = 40 \times 10^3$ for a helix of infinite length, k the wavelength dependent constant at 222 nm = 4.6, and N is the number of residues in the peptide chain . The number of helical residues was calculated by multiplying the % helical content by the total number of residues in the peptide.

d The ratio of the mean residue molar ellipticities of 222 nm to 208 nm.

e $[\mathrm{GdnHCl}]_{1/2}$ is the transition midpoint, the concentration of guanidine hydrochloride (M) required to give a 50% decrease in molar ellipticity at 222 nm between the folded and unfolded states.

f m is the slope in the equation $\Delta G_u = \Delta G_u^{H_2O} - m[\mathrm{GdnHCl}]$ (Pace *et.* [illegible]. 1988).

g $\Delta\Delta G_u$ is the difference in free energy of unfolding relative to peptide K3b. $\Delta\Delta G_u$ was calculated using the equation:
$\Delta\Delta G_u = ([\mathrm{denaturant}]_{1/2,A} - [\mathrm{denaturant}]_{1/2,B})(m_A + m_B)/2$ (Serano and Fersht, 1989); see methods for more detail. The $\Delta G_u^{H_2O}$ for peptide K3b was 9.41 kcal/mol.

* Peptides studied previously (Tripet *et al.* 1997).

resulted in a significant stability decrease of the coiled-coil region [compare GdnHCl midpoints of K1b (0.82 M) and K4 (3.73 M); Fig. V-3B and Table V-1.]. Thus, it would appear that residues within the last heptad, although not required for helix induction *per se*, are absolutely required for the high association and stability of the kinesin necks dimeric structure.

To define clearly the location of the N-terminus of the coiled-coil dimerization domain in solution [i.e, to confirm its predicted initiation at residue 337 (Fig. V-1) as well as its observed initiation at this residue within the crystal structures of rat brain kinesin (Kozielski, *et al.*, 1997; Sack, *et al.*, 1997)], we prepared and analyzed the kinesin peptide variant K3b (sequence shown in Fig.V-2). This peptide corresponds to residues 337-383 of the human kinesin neck region which encompasses all 5 of the predicted coiled-coil heptads in addition to the 15 residues to its C-terminus. CD analysis of the K3b peptide (Fig. V-3C) shows a highly α-helical spectrum which is greater in mean residue molar ellipticity than both kinesin peptide K3 (residues 344-383), which lacks the first 7 residues (first heptad) of the coiled-coil domain, and K5 (330-383), which contains the first heptad but is further extended by seven residues towards the N-terminus. The molar ellipticity value, (-23, 598°; Table V-1) observed for this peptide (K3b) corresponds to a calculated α-helical content of 66% or ~31 helical residues, thus closely matching the observed number of α-helical residues within the rat brain dimer crystal structure coiled-coil (32 residues) (Kozielski, *et al.*, 1997) as well as that predicted by our prediction method (33 residues). Thus, it would appear that all of the α-helical residues within the neck region are indeed located entirely within the region from residue 337 to 370.

It is noteworthy that the significantly greater degree of molar ellipticity observed for the K3b peptide (-23,598°) in comparison to K5 (-19, 500°) clearly demonstrates that the 7 beta-linker residues (330-336) located N-terminal to the coiled-coil region are structurally "non-helical", suggesting that these residues have within them the intrinsic ability to terminate the α-helix structure at residue 337 and adopt an alternative secondary structure independent of other interactions.

To assess the contribution of the first heptad towards the stability of the coiled-coil domain, the stability of the K3b peptide was next determined. The GdnHCl denaturation profile (Fig. V-3D) showed a midpoint of 2.93 M and an extrapolated ΔG_u of folding of 9.41 *kcal·mol*$^{-1}$ (Table V-1). As expected, based on its greater helicity, the K3b peptide was significantly more stable than that of the K3 peptide which lacked the first heptad (compare GdnHCl midpoints of 1.17 M and 2.93 M, resectively; Table V-1), however, quite suprisingly, the K3b peptide was not as stable as that of the K5 peptide which contained the entire five heptads of the neck coiled-coil in addition to the 7 N-terminal beta-linker residues (compare GdnHCl midpoints of 2.93 M and 3.93 M, respectively; Table V-1) . This significant difference (~1 M GdnHCl) would indicate that although the 7 residues extending from the N-terminus of the coiled-coil are structurally "non-helical", they are still interacting in a positive manner to stabilize further the coiled-coil structure.

Kinesin neck coiled-coil interactions with the N-terminal β-linker region.

To investigate further how residues from within the β-linker region are stabilizing the coiled-coil structure, we investigated whether this interaction is influenced by the stability of the coiled-coil and/or characteristics of specific residues within the coiled-coil ? That is, we turned our attention to the region encompassing residues 344-351, being noteworthy for containing several "non-ideal" hydrophobic core residues as well as exhibiting the large dip in the predicted stability profile of the coiled-coil region (Fig. V-1B). Thus, we prepared the two kinesin peptide variants K2b and K2c (sequences shown in Fig. V-2) which encompass regions 337-369 and 330-369, respectively, but which contained Lys 343, Glu 347, and Lys 350 replaced with Glu 343, Leu 347 and Glu 350, respectively. These substitutions designed to remove the destabilizing effect of the glutamic acid (Glu 347) within the hydrophobic core position and introduced favourable (i-i'+5) electrostatic

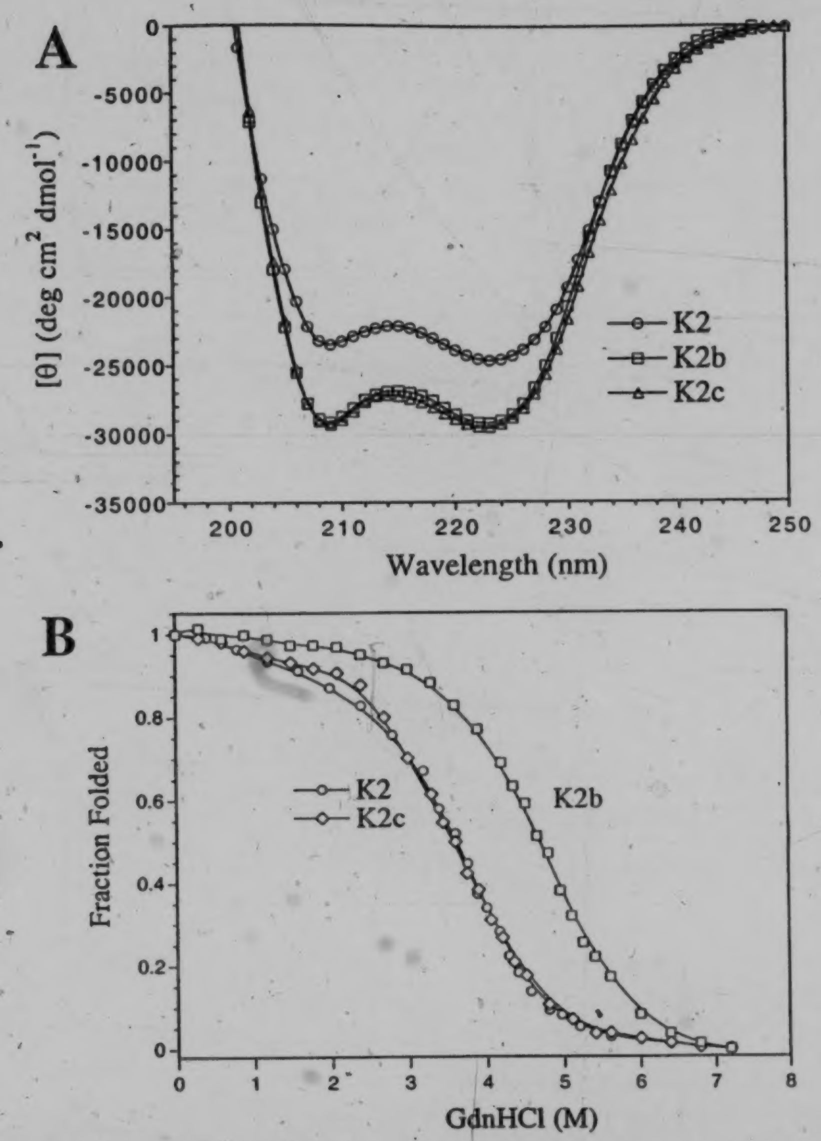


Figure V-4. *Panel A*; CD spectra of kinesin peptides K2, K2b, and K2c. Spectra were recorded in a 0.1 M KCl, 0.05 M PO_4, 0.002M DTT, pH 7 buffer. All peptide concentrations were 100 µM. ***Panel B***, GdnHCl denaturation profiles of kinesin peptides K2, K2b and K2c. Denaturation profiles were recorded at 222 nm at 20°C in a 0.1 M KCl, 0.05 M PO_4, 0.002 M DTT , pH 7 buffer with GdnHCl as the denaturant. The fraction folded (f*f*) was calculated as that described in Figure V-3. Peptide concentrations were 60 µM. CD spectra and denaturation profiles for kinesin peptides K2 have been reported previously (Tripet *et. al.* 1997) but are shown for direct comparison to the two new kinesin peptides K2b and K2c.

interactions across the coiled-coil interface (Glu 343 with Lys 348'and Glu 350 with Arg 355'). As a preliminary test (prior to synthesis) to verify the choice of substitutions would indeed cause the desired increase in coiled-coil stability, we re-analyzed the kinesin neck sequence using the above predictive method and new sequences to generate a new stability profile (Fig. V-1B; the difference in stability between the native and mutated sequence is highlighted by shading) . As expected, the region of "inherent low stability" was now predicted to be absent.

Figure V-4A shows the CD analysis of kinesin peptides K2b and K2c. In agreement with the prediction method, both K2b and K2c showed significantly greater degrees in mean molar ellipticity relative to the K2 peptide (Fig. V-4A and Table V-1). Comparison of the CD spectra between K2b (residues 337-369) and K2c (residues 330-369) showed that the addition of the 7 N-terminal β-linker residues (330-336) again did not enhance the molar ellipticity of K2b relative to K2c. However, calculation of the number of helical residues for the two peptides did show that upon addition of the 7 N-terminal residues, the number of α-helical residues increased from 28 (K2c) to 33 (K2b). Thus, although the β-linker residues are 'non-helical' they are again interacting in a positive manner to increase further the helicity of the N-terminus of the coiled-coil structure.

To investigate whether the β-linker residues still impart the significant effect on the stability of the coiled-coil domain as observed between K3b and K5, K2b and K2c were denatured by GdnHCl. As shown in Fig. V-4B and Table V-1, the stabilizing effects of the designed substitutions is clearly reflected by the greater stability of the K2b and K2c peptides (transition midpoint of 4.75 M and 3.6 M, respectively) relative to K3b (2.93 M) . Note that addition of the 7 N-terminal residues again caused an increase in the GdnHCl transition midpoint (from 3.6 M for K2c to 4.75 M for K2b; ΔGdnHCl midpoint of 1.15 M) comparable to that observed above between the native kinesin sequence peptides K3b (2.93 M) and K5 (3.93 M) (ΔGdnHCl midpoint = 1.0 M; Fig. V-3D and Table V-1). Thus, the N-terminal β-linker stabilizing interactions appear to be independent of the stability of the coiled-coil structure or any potential effects of the surface exposed charged lysine residues (343 and 350).

Effect of helix-capping interactions between the β-linker and the coiled-coil regions on the stability of the kinesin neck coiled-coil.

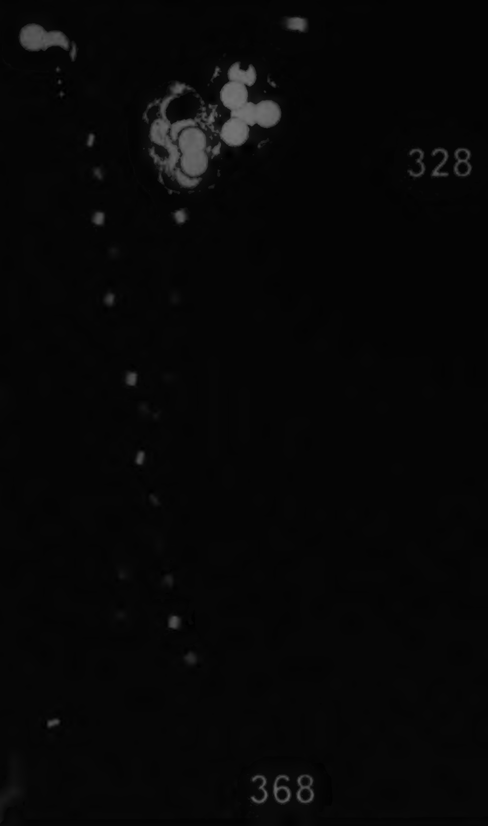
To reconcile how residues from within the β-linker region are imparting such a significant effect upon the neck region coiled-coil's stability and helicity, we analyzed the recently solved X-ray crystallographic structure of dimeric rat brain kinesin (Kozielski, *et al.*, 1997) for the appearance of any unique structural features which exist specifically between residues within the β-linker and the coiled-coil domain . Interestingly, inspection of the β-linker coiled-coil interface showed that the sidechain of Leu 335 (in human sequence numbers) within the β-linker was interacting with the exposed surface of the indole ring of Trp 340 which was not shielded within the hydrophobic core of the coiled-coil. Additionally, the side chain of Thr 336 (within the β-linker) was close to Gln 339 (within the coiled-coil). Although not specifically stated by the authors, these two side chain interactions depict the two classic helix capping motifs: the N-cap and the hydrophobic staple (Munoz, *et al.*, 1995; Richardson and Richardson, 1988; Seale, *et al.*, 1994). The N-cap arises from the side chain interactions of the Thr 336 hydroxyl group with the backbone amide of residue Gln 339 (Fig. V-5), and the hydrophobic staple arises from the i to i+5 hydrophobic interactions between Leu 335 and Trp 340 within the coiled-coil region (Fig. V-5). Moreover, Munoz *et al.*, (1995) have recently shown that such interactions can significantly stabilize α-helices. Further, the synergistic interaction of both motifs can significantly stabilize such structures. Thus, it would seem plausible that it is these specific structural interactions between the β-linker region and the N-terminus of the coiled-coil domain which are responsible for the observed increases in stability.

To establish if this is indeed the case, we analyzed the secondary structure content and stability of the two kinesin peptide variants K2e and K2f (sequences shown in Fig. V-2) in

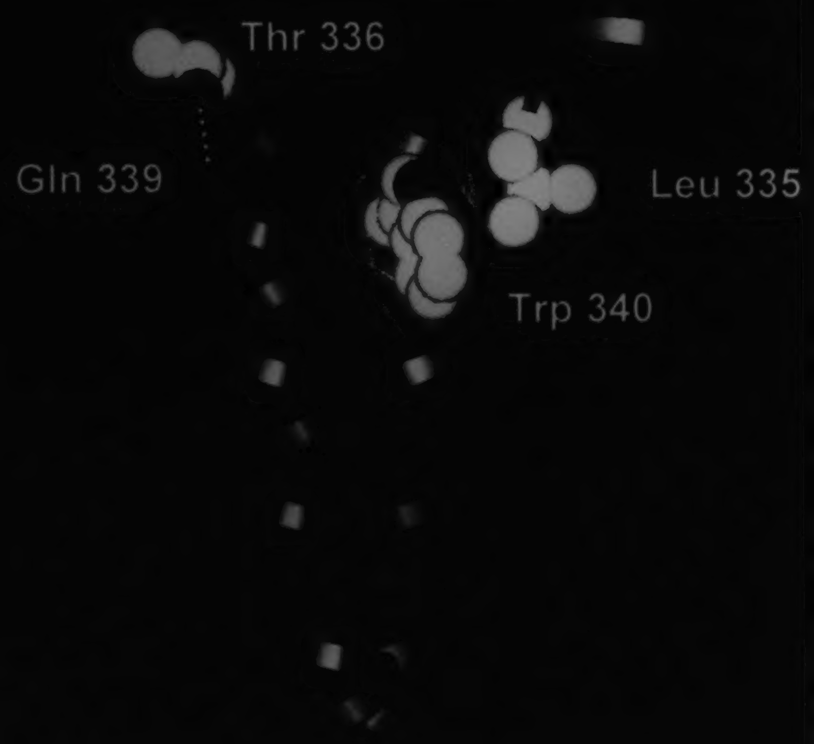
---

Figure V-5. Molecular interactions occurring between Leu 335, Thr 336, Glu 339 and Trp 340 at the N-terminus of the rat brain kinesin neck. The X-ray crystallographic coordinates of the dimeric rat brain kinesin were used (Kozielski *et. al.* 1997). Note amino acid residue numbers are shown in reference to the human kinesin sequence, in rat brain kinesin the amino acid residue numbers are Leu 337, Thr 338 , Glu 341 and Trp 342, respectively). In solid grey: ribbon representation of the kinesin neck region from residues 328 to 368. The position of Leu 335, Thr 336, Glu 339 and Trp 340 are labeled. Left side of the coiled-coil structure depicts the N-cap effect of Thr 336 hydroxyl (red sphere) with the amide nitrogen (blue sphere) of the helix backbone residue Glu 339, Right side of the coiled-coil structure depicts the hydrophobic staple motif occuring between Leu 335 (white spheres) with Trp 340 (yellow spheres). Hydrophobic surface packing interactions are emphasized by the van der waal surface interaction (dotted sheres).

Kinesin Neck region



N - Cap and Hydrophobic Staple


Thr 336
Gln 339
Leu 335
Trp 340

which we systematically substituted each of the helix capping residues (Leu 335 and Thr 336) with glycine residues. Far UV CD spectroscopy analysis of K2e (Leu 335 Gly) demonstrated only a small decrease in its α-helical content as compared to the K2b peptide, indicating little effect upon the α-helical residues within the coiled-coil region (Table V-1). In contrast, stability analysis demonstrated a large decrease in its GdnHCl denaturation midpoint value relative to K2b (compare GdnHCl midpoints 3.71 M and 4.74 M respectively). Substitution of Thr 336 with Gly in addition to a Leu to Gly 335 substitution (peptide K2f) resulted in a further decrease in peptide helical content and stability (compare midpoint values of 3.35 M and 3.71 M for K2f and K2e, respectively), however, the latter decrease was significantly less (only 0.36 M) than that observed for the Leu substitution (1.04 M). Interestingly, the stability of K2f (3.35 M), with both helix capping residues now substituted with Gly residues, was comparable to that of K2c (3.60 M) which contained no β-linker residues. Thus, the helix capping residues Leu 335 and Thr 336 from within the β-linker region are responsible for the increased stability of the coiled-coil region in the presence of the linker region.

Examination of the potential role of Trp 340 in unfolding the coiled-coil domain.

In light of the above observation which showed the first heptad to play such a critical role in stabilizing the neck region coiled-coil (compare K3 with K3b; Fig. V-3D), combined with the observation that a hydrophobic staple motif occurs directly between residues of Leu 335 and Trp 340 (which specifically resides within the first heptad), has thus raised the question whether modulation of the hydrophobic contribution of Trp 340 toward the hydrophobic core of the first heptad of the coiled-coil results in opening of the coiled-coil domain, i.e., is Trp 340 the critical trigger residue in folding and unfolding the N-terminus of the kinesin coiled-coil domain ?

To investigate this hypothesis, we analyzed the secondary structure content and stability of the two kinesin peptide variants K2d and K2g (sequences shown in Fig. V-2) which

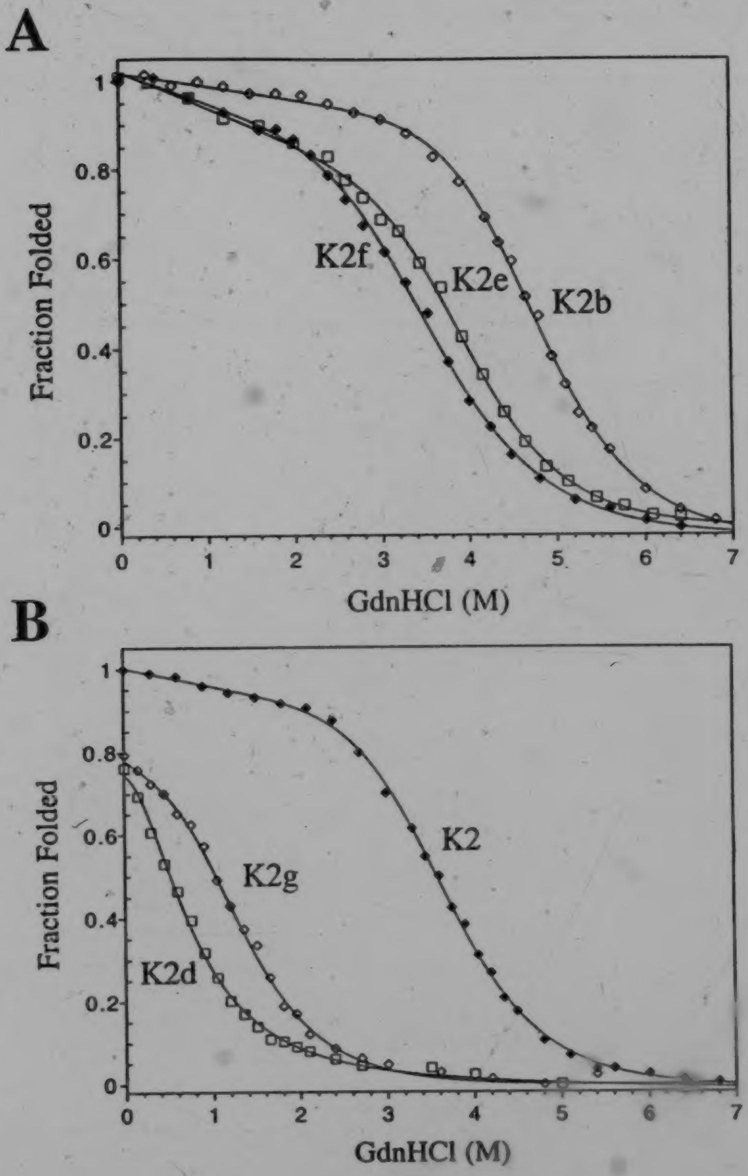


Figure V-6. Substitution of Leu 335, Thr 336 and Trp 340 residues. *Panel A*; GdnHCl denaturation profiles of kinesin peptides K2b, K2e and K2f. Denaturation profiles were recorded at 222 nm at 20°C in a 0.1 M KCl, 0.05 M PO_4, 0.002 M DTT , pH 7 buffer with GdnHCl as the denaturant. The fraction folded (f_f) was calculated as that described in Figure V-3. Peptide concentrations were 100 μM. *Panel B*, GdnHCl denaturation profiles of kinesin peptides K2c, K2d and K2g. Denaturation profiles were recorded as above. Peptide concentration were 100 μM.

contained Trp 340 substituted by Ala, a residue of high helical propensity but of low apparent hydrophobicity (Monera, *et al.*, 1995). Far UV CD analysis of the K2g peptide demonstrated only a small decrease in molar ellipticity in comparison to K2 (-23,111° and -24,600°, respectively; Table V-1), thus indicating largely no loss in the α-helical content of the coiled-coil domain. In contrast, stability analysis of K2g demonstrated a significantly lower GdnHCl midpoint value (compare 1.17 M for K2g versus 3.61 M for K2, Fig. V-6 and Table V-1). Thus although not required for helicity *per se*, the interaction of the Trp side-chain is absolutely critical for the stability of the coiled-coil domain. It is interesting to note that the decrease in stability (ΔGdnHCl midpoint of 2.44 M) is greater than the loss of the hydrophobic stable motif by the Leu 335 to Gly substitution (compare K2b (4.75 M) and K2e (3.71 M) with ΔGdnHCl midpoint of 1.04 M) and is analogous to the loss in stability observed when the β-linker region and the first heptad was completely removed [compare stability values of 3.93 M (K5) to 1.17 M (K3); ΔGdnHCl midpoint of 2.76 M].

Analysis of the K2d peptide (which lacked the 7 N-terminal β-linker residues) showed that this large decrease in stability was completely independent of the 7 N-terminal beta-sheet residues (compare K2d (0.5 M) to K3b (2.93 M) for a ΔGdnHCl midpoint of 2.43 M, Table V-1). Note that the stability difference between K2d and K2g (ΔGdnHCl midpoint of 0.67 M) is within experimental error of the loss of the N-cap interaction (compare ΔGdnHCl of 0.36 M for K2e to K2f).

Thus taken together, these results indicate that Trp 340 does play a critical role in stabilizing the first heptad of the coiled-coil, and modulation of its hydrophobic contribution towards the hydrophobic core possibly via a change in the helix capping interaction of Leu 335 and Trp 340, could indeed be a triggering mechanism for reducing the stability of the N-terminus of the neck coiled-coil to allow unfolding of the neck coiled-coil region during the mechanical stepping of the kinesin motor protein (see model below).

C. Discussion

In an effort to understand better the mechanism by which dimeric kinesin is able to move processively along microtubules, we have investigated the secondary structure content and stability of several new synthetic peptide variants corresponding to the human kinesin neck region. Our results show that the neck region coiled-coil domain exists from residues 337 to 369, as assessed by solution CD studies, and thus is in excellent agreement with that of the recently solved X-ray crystallographic structures of monomeric and dimeric rat brain kinesin (Kozielski, *et al.*, 1997; Sack, *et al.*, 1997; Thormahlen, *et al.*, 1998). These structural studies show the neck region helix to begin at Ala 337 and proceed to ~ Trp 368. Moreover, the secondary structural studies have shown that the definition of the neck region appears to be the result of distict structural features within its sequence e.g. the two helix capping residues located at its N-terminus which can further modulate the stability of the coiled-coil and thus be involved in its function (see below).

Importance of the N and C terminal heptads of the kinesin neck coiled-coil.

Analysis of the stability of two new kinesin peptide variants, K1b and K3b, has further demonstrated that full stability of the coiled-coil region is critically dependent on the presence of all five of the coiled-coil heptads, with the first and last heptad playing particularly critical roles. For example, GdnHCl denaturation studies of the K1b peptide demonstrated that deletion of the last 6 residues from the last heptad results in a significant loss in the observed stability of the coiled-coil domain (compare $[\mathrm{GdnHCl}]_{1/2}$ values of 3.73 M for K4 versus 0.82 M for K1b). The degree of molar ellipticity and 222/208 nm ratio for K1b indicate, that although this region can still adopt an α-helical conformation, it is now no longer able to form a highly associated dimeric structure. These results appear to be in excellent agreement with previously characterized recombinantly expressed *Drosophila*, Rat and Human kinesin motor domain constructs which have demonstrated that truncation of any of the residues within the last two C-terminal heptads results in motors domain constructs which remain largely monomeric in their oligomerization state (Correia, *et al.*, 1995; Huang, *et al.*, 1994; Jiang, *et al.*, 1997; Sack, *et al.*, 1997)

As noted above, the last two heptads of the kinesin neck coiled-coil demonstrate sequence homology with a sequence that has been implicated in nucleating coiled-coil interactions (i.e., a 'trigger sequence'). Based on the above results, which clearly show the continued presence of α-helical conformation upon truncation of the C-terminal 6 residues, combined with those of Morii *et al.* (1997) and Sack *et al.* (1997) which have also shown the presence of α-helical structures in the neck regions in the absence of both heptads, we would have to conclude that the requirement of the last two heptads for induction of the helical structure or to keep the N-terminus closed does not appear to be correct. However, the last two heptads do contribute substantially to coiled-coil stability.

The significance of the first heptad of the coiled-coil neck region was rather surprising. As noted above (Fig. V-1), the first heptad (residues 337-343) is not predicted to contribute greatly to the stability of the coiled-coil region. Thus, initially we rationalized its significant contribution must be a result of 'non-classical' coiled-coil interactions. That is, Previously (Tripet, *et al.*, 1997), we noted that the apparent high stability of the neck coiled-coil region, despite the presence of a glutamic acid residue (Glu 347) within a **d** position of the hydrophobic core must arise as a result of the four lysine residues located in the adjacent **e** and **g** positions, two of which specifically arise from within the first heptad (Lys 341 and Lys 343). Hence, the significant loss in stability upon the deletion of the first heptad was believed to be a result of the loss of both the stabilizing electrostatic interactions which were mitigating the destabilizing effect of Glu 347 as well as the loss of the contribution to the hydrophobic core of the methylene groups of the Lys side-chains. More recently, however, analysis of the crystal structure of the neck region coiled-coil by Madelkow's group (Thormahlen, et al., 1998) showed that although the lysine side chains are located adjacent to the hydrophobic core and close to Glu 347, no direct salt bridge interactions were apparent. Furthermore, they suggested that Glu 347, together with Tyr 344, was possibly stabilizing the coiled-coil by not allowing the lysine sidechains to interact directly across the core, thus preventing a repulsive effect (Thormahlen, *et al.*, 1998). In agreement with these findings, investigation into this region using the K2d and K2g peptide variants (which contained Trp 340 substituted by Ala) showed that substitution of the single large hydrophobic residue of Trp 340 is sufficient to allow opening of the coiled-coil despite the presence of the charged lysine residues . Thus, the stability of the first heptad appears to be the result of unique structural interactions occurring exterior to the hydrophobic core.

Taken together, it would appear that the stability of the kinesin neck coiled-coil is more representative of a finely tuned 'spring' such that the weak but significant hydrophobic interactions within the first heptad of the N-terminus can generate an extremely stable coiled-coil, however, if modulated (i.e. the first heptad) possibly via a structural change within the N-terminal β-linker region via the helix capping residues (see below), could result in its opening.

Beta-linker and coiled-coil interactions

Implicit in the design plan of a coiled-coil that can unwind and rewind is the presence of a mechanism by which the conformational status of the individual motor domain heads can be communicated to the neck region coiled-coil. Analysis of the kinesin peptides K3b, K5, K2b and K2c which span this region indicated that residues from within the beta-linker and the coiled-coil are structurally linked and thus communication is possible. That is, comparison of the stability in both cases (ie., K2b vs. K2c and K3b vs K5) indicates that residues from the beta-linker region can further stabilize the coiled-coil by as much as ~ 1 M GdnHCl or ~1.5 kcal mol^{-1}. Specific substitutions within the linker region have indicated that the enhanced stability arises primarily from the side chain residue interaction between Leu 335 and the sidechain of Trp 340, in agreement with that observed in the x-ray crystallographic structure of dimeric rat brain kinesin. Although not specifically stated by the authors, such side chain residue interactions are representative of the hydrophobic staple helix capping motif (Munoz, *et al.*, 1995). In this motif, the residue of the N' position makes packing interactions with the N4 residue of the helix. Such interactions result in very sharp turns of the loop region as observed in the kinesin structure. To the best of our knowledge, no such interactions have been cited as specifically occurring at the top of a two-stranded coiled-coil structure, although such interactions appear to be prevalent in α-helices in proteins (Seale, *et al.*, 1994). Thus, it appears that kinesin has evolved a novel way to utilize such a motif at the ends of its neck coiled-coil.

Munoz *et. al.* (1995) have investigated the stability effects of the hydrophobic staple and shown that the presence of both helix capping motifs, i.e., the helix N-cap as well as the hydrophobic staple, can act synergistically to produce an autonomous folded unit in solution. Thus, kinesin has efficiently optimized the full use of such structural features important for helix capping. It is important to note that, although such motifs make helices more energetically stable, more importantly they are implimented for structural demarcation of the end of the helix. In the case of the kinesin neck region, this may be important to make sure helical induction is not transfered to the beta-linker (beta sheet secondary structure) region or vice versa. Conversely however, because they can contribute energetically to the stability as observed above (Fig. V-3 and V-4), modulation of such interactions can thus potentially regulate the stability of the coiled-coil. Moreover, modulation of such interactions could conceivably regulate the structure within the neck region, i.e., movement of Leu 335 could now allow further helical induction. In either scenario, the more important consequence of the observation of the presence of this motif is that it could potentially play a key role in regulating the folding and unfolding process of the neck coiled-coil during mechanical stepping. From our analysis, we perceived that the major energetic barrier towards unfolding of the coiled-coil (at least the first heptad) would be exposure of the hydrophobic surface area buried within the core **a** and **d** positions. In the case of kinesin, this would be largely the exposure of Trp 340. The presence of the hydrophobic stable motif directly involved with this residue thus allows a mechanism by which such exposure could occur. That is, a conformational change within the beta-linker region driven by the motor domain could potentially drive the partial re-orientation of Leu 335 resulting in a further increase in hydrophobic interactions and van der Waals contacts with Trp 340. This would likely maintain the presence of the individual helices of the first heptad but now let each of the individual helices, together with the second heptad to separate transiently to carry out the next 8 nm step. In support of such a proposal is the observation that loss of the hydrophobic packing of this residue (Trp 340 ->Ala 340) resulted in no loss of helical content but significant loss in the stability, similar to that observed for the deletion of the first heptad from the coiled-coil region (peptide K3), thus demonstrating the importance of the Trp residues interaction in maintaining the N-terminus of the coiled-coil and its ability to switch between a fully helical state and the transient unfolded state.

Further support for such a triggering mechanism arises from the recent high resolution structural studies from Mandelkow's group (Kozielski, *et al.*, 1997; Sack, *et al.*, 1997; Thormahlen, *et al.*, 1998) who showed the greatest divergence between the monomeric and dimeric rat brain structures occurs at Asn 334 (332 in the human sequence) at the beginning of beta strand 10 (β10). This divergence is accompanied by related changes at the bottom of beta-strands β7 and β6 to which β10 is connected by H-bonds, suggesting that conformational changes within the motor domain (i.e., at switch I and switch II) could be easily transmitted to the Leu 335 region. Moreover, Sack *et al.*, (1997) have also noted that the connection to the bulk of the head via β9 and β10 "appears quite loose and that it is conceivable that a change in direction might take place at the β10-α7 junction during the ATPase cycle or powerstroke" (Sack, et al., 1997), supporting a significant change in the region of Leu 335.

Interestingly, in the monomeric structure, Leu 337 (β10) shows very close packing contacts to Lys 222 of loop 10 of the core of the motor domain, indicating that the packing interactions can be broken or switched depending upon the coiled-coil dimerization state.

In fact, as noted by Thormahlen *et al.*, (1998), it appears that Trp 342 (340 in the human sequence) clearly changes orientation upon dimerization. Hence, head to head communication between the two motor domains could easily be communicated via this mechanism. Additionally, the helix capping residue interactions between Leu 335 and Trp 340 are also likely to represent an important structural link required between the β-linker region and coiled-coil region, allowing for the specific directional biasing of the free unbound kinesin motor domain head towards the plus end of the microtubule.

Does the kinesin neck region coiled-coil open or not ?

Although the above structural data appear to be in good agreement with an unfolding scenario for the adjacent coiled-coil region in kinesin processivity, recent functional data surrounding this issue do not appear to be entirely self consistent. For instance, several strong arguments against such an unfolding scenario of the coiled-coil region have come from the recent results of Romberg *et al.*, (1998), who both mutated the region of inherent low stability within the coiled-coil region with residues of high stability, as well as replaced the native neck region coiled-coil of human kinesin from residue 344 to 372 with that of an exteremely stable *de novo* designed coiled-coil domain (STABLECOIL). Both of these mutations caused only small decreases in the level of processivity. Thus, significant unwinding of the coiled-coil does not appear to be an absolute requirement for processivity. However, the results of these authors do not rule out that such unwinding may also come from the first heptad of the coiled-coil region, as this region was left intact in both mutated constructs. In fact, some of their results do support partial unfolding/separation of the first heptad of the coiled-coil. For example, strengthening the coiled-coil domain by substitution of the "non-ideal" coiled-coil residues to more "ideal " residues did in fact decrease processive movement, albeit only marginally (2-fold). As these mutations reside structurally only inside the dimerization interface, and are only expected to stabilize the coiled-coil region further, a small amount of separation between the two strands of the coiled-coil within the coiled-coil region must therefore be important. Furthermore, Romberg *et al.*, (1998) have shown that duplication of the first heptad resulted in a construct which appeared to be more processive (3-fold). If unwinding (or separation) of the first heptad is important for proceeding to the next microtubule binding site, then duplication of this region would result in a region which could unwind and extend further, thus positioning the forward motor domain head either closer to the next microtubule binding site more efficiently, or even allowing the forward positioned head to bind the next adjacent microtubule binding site. The latter proposal would seem to agree more closely with their results which also showed a lower microtubule binding constant for the DUP construct. Although Romberg *et al.*, (1998) did allude to such a possibilities, they suggested this may not be the case based on the fact that a second construct which contained a three glycine residue insert in the same location produced the opposite effect, i.e., decreased activity. Based on our present results, however, which clearly demonstrate that the interactions of the first heptad of the coiled-coil and β-linker region play such a critical role in the overall stability of the kinesin neck coiled-coil dimerization domain, the insert of three glycine residues, which are noted to be strong α-helical breakers, would in essence essentially terminate the coiled-coil domain 7-residues earlier (at residue 344) and by comparison to the synthetic peptide K3, which lacked the first heptad, may have resulted in a construct which now lacked sufficient stability within the neck region to form a coiled-coil or even maintain dimerization. Thus the loss in processiveness of this construct may in fact be a result of the overall loss in dimerization stability within the neck region even though the linker region has been extended. Additionally, the loss in activity due to the Gly spacer may highlight the need for the neck region to adopt both a helical conformation and an extended conformation for efficient movement.

By far the strongest aruguements against unwinding of the coiled-coil arise from the observation that several of the mutant constructs, particularly those having a stabilized coiled-coil domain, continue to remain attached to the microtubule filament for extended periods of time. We would like to emphasize, however, that caution should be taken in interpreting these results as an absolute test of "native" motor processivity. Native kinesin is normally attached to a cargo domain and thus must work continually in the context of a physical load, which may resist its movement or detach it as it travels along the microtubule filament. For this reason, analogous native-like motor processivity should also be tested under such constraints or conditions (i.e., to do work as in the optical trap experiments). This would then verify the ability of each of these new motor domain constructs to reach the next microtubule binding site, an absolute requirement in the native setting. It is interesting to note that preliminary optical trap experiments by Romberg *et al.*, (1998) observed the complete absence of any processivity (personal communication, R. D. Vale 1998) for of the stablized kinesin motor domain constructs. Thus, we rationalize their reported results (Romberg *et al.*, 1998) as representing the ability of the motor molecules to still be able to extend the leading motor domain head towards the next microtubule binding site to a point where it can "almost bind" the next site. In the absence of a physical load, this structural biasing is sufficient statistically to allow the motors to remain bound to the filament as opposed to diffusing away. In the presence of a load, however, the inability of the motors to efficiently extend the neck coiled-coil region and therefore 'absoultely' reach the next microtubule binding site stalls the motors and/or removes them from the filament.

Is helix capping a unique structural feature within conventional kinesins only ?

The observation of specific helix capping residues within the N-terminus of conventional kinesins raises the interesting question of whether such structural features are also present within the other N-type kinesin-like motor types. Preliminary sequence analysis of these neck regions (Chana *et al.* 1999, manuscript in preparation) has shown the absence of the conservation of these specific residues, however, the specific 'N1'-helix capping residue proline is observed in the same location as Leu 335. Thus, termination of the helical neck region directly at the begining of β-10 is present, as well as a mechanism to

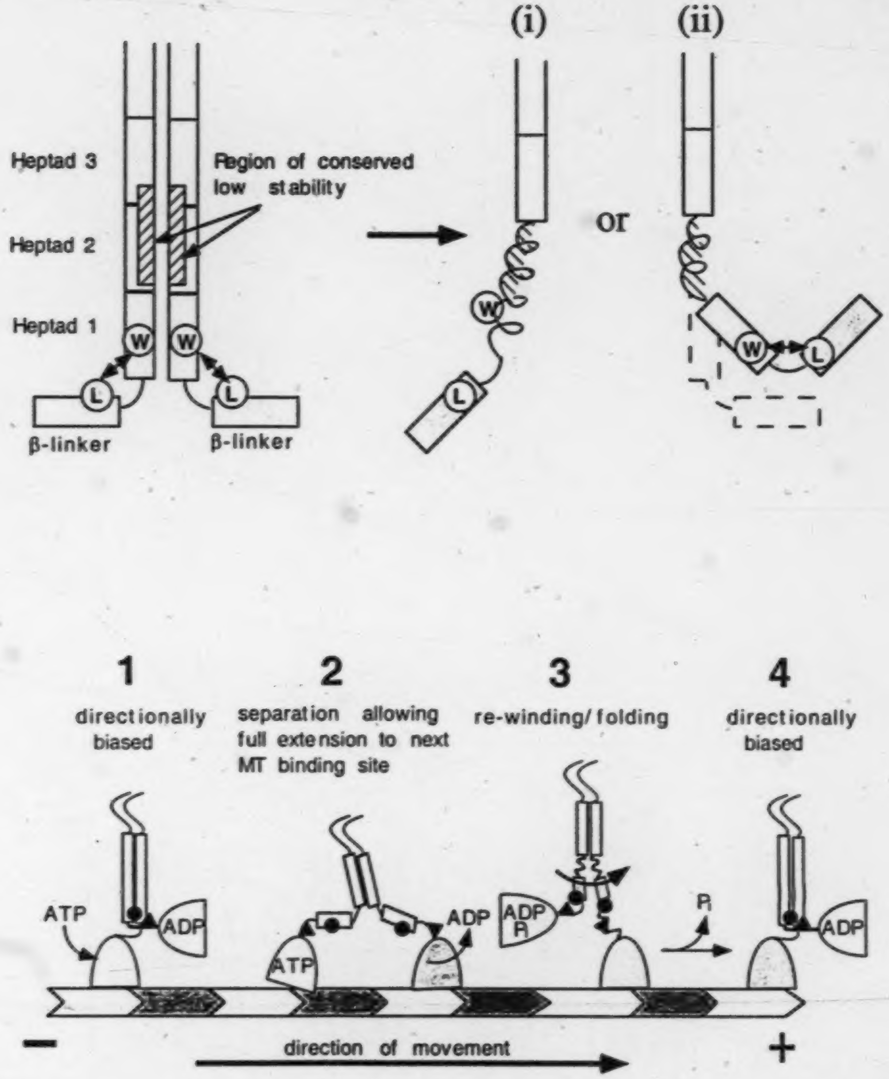

(i)
(ii)
Heptad 3
Heptad 2
Heptad 1
Region of conserved low stability
W
L
β-linker
β-linker
or
1
directionally biased
2
separation allowing full extension to next MT binding site
3
re-winding/folding
4
directionally biased
ATP
ADP
ADP
ADP Pi
Pi
ADP
−
direction of movement
+

Figure V-7. Top) Organization and possible senarios for the helix capping residues in processivity. Location of the Leu-Trp capping interaction in relation to the conserved region of low stability of the coiled-coil is shown. Two possible senarios for the helix capping residues include (i) β-linker region changes conformation to cause loss of the end capping effect resulting in a decrease in stability which then allows for the coiled-coil to open up, (ii) β-linker region changes conformation to cause Leu 335 to pull Trp 340 away from the hydrophobic core resulting in the opening of the coiled-coil domain due to its lost interactions. **Bottom)** Model for kinesin processivity. *Bell shaped* spheres depict the motor domain heads, *rectangular boxes* indicate the kinesin neck region which forms a two stranded α-helical coiled-coil, the *shaded rectangles* at the base indicate the α and β tubulin subunits. For a more detailed description of each step (A to D) in the model see "Discussion" .

create a specific 90° turn. Moreover, preliminary analysis of the stability of these neck regions in the presence or absence of the linker region has shown no increase in the stability of the coiled-coil regions, suggesting the absence of the type of communicative link observed within conventional kinesins. Thus, the absence of these features combined with the absence of the "destabilizing residues" within the coiled-coil region may be one factor as to why these motors may not be processive.

A new model for kinesin processivity.

Several models describing the mechanism of kinesin processivity have emphasized the role of the "destabilizing residues" within the coiled-coil as being important for allowing unfolding of the neck region during stepping to the next microtubule binding site (Cross, 1995; Hackney, 1994; Hirose, *et al.*, 1995; Hirose, *et al.*, 1996; Ma and Taylor, 1997; Moyer, *et al.*, 1996; Tripet, *et al.*, 1997). Although the results of Romberg *et al.*, (1998) initially appear to contradict such a proposal, suggesting unfolding primarily occurs from within the β-linker region itself, their results, in conjuction with those of the present study, may in fact be rationalized into a stronger model for processivity. Thus, unfolding could be occuring from within both the beta-linker region and the first two heptads of the coiled-coil, being directly communicated via a conformational change between the β-linker and the helix capping residues of Leu 335 and Thr 336. Such a hypothetical model is described below in the context of other recent structural and functional studies on kinesin. In this model (Fig. V-7), a dimeric kinesin molecule is initially bound to the microtubule filament anchored by one of its motor domain heads (without nucleotide) in a strong microtubule binding state, while its second head domain (with ADP bound) is detached and pointing towards the plus end of the microtubule. In this state (Step 1), direction of the detached motor domain head is oriented and stabilized by intramolecular interactions occuring between the beta-linker region-helix capping residues and the N-terminus of the coiled-coil domain. Similar intramolecular interactions are also present in the bound motor domain head. In this state, the neck beta linker region of the detached motor domain head is proposed to exist in a "cocked" beta-sheet-like structure as observed in the dimeric crystal structure (Kozielski, *et al.*, 1997). The 5 heptads of the neck domain are also proposed to exist in a coiled-coil structure which constrains the detached head from reaching the next microtubule binding site.

Upon binding ATP, the bound motor domain undergoes a conformational change which changes the interactions of the β1-β9-β10 interaction and the helix capping interactions between Leu 335 and Trp 340. This change in interactions is envisioned to be either a re-orientation of Leu 335 back to the motor domain head (Fig. V-7, (i)) resulting in a complete loss in the Leu 335-Trp 340 interaction, or a torque which opens the helix capped first heptad (Fig. V-7 (ii)) . The resultant change in interactions causes a transient loss of stability in the coiled-coil region which results in the separation of the first two heptads (in analogy with our Trp 340 to Ala substitution). The concomitant conformational changes within the neck region coiled-coil beta-linker region also release the detached head from its "cocked" state to one that now re-orientates the beta-sheet region, as proposed by Romberg *et al.*, (1998), allowing full extension to the next microtubule binding site. The transient loss in the coiled-coil and beta-linker is viewed as a high energy mechanical state (tensioned spring). Upon binding to the next microtubule binding site (Step 2), the leading motor domain releases its bound ADP and switches into a strong microtubule binding state/conformation (Crevel, *et al.*, 1996; Hackney, 1994; Hackney, 1994). This change in conformation driving the beta-linker region to return back into its β-sheet conformation, initiating contraction and thus causing a forward force. Based on the studies of Hancock and Howard (Hancock and Howard, 1998), this change would be responsible in part for the strain or pull required to accelerate the release of the trailing head. In step 3, the released lagging motor domain head concurrently hydrolyzes ATP to ADP, thus switching its conformational state (Hirose, *et al.*, 1995), and re-initiating or returning the helix capping interactions between the beta-linker region and the coiled-coil to its coiled-coil forming state. The general restructuring of the beta-linker regions (back into a sheet secondary conformation) allows for the first two heptads of the coiled-coil to "re-zip" and thus re-orientates the released motor domain head ~180°, towards the neck microtubule binding site, step 4.

In the present model, the coiled-coil acts as one part of the transducer by (1) momentarily extending the distance between the two motor domain heads during stepping; and (2) acting as the spring to help drive the fast repositioning of the released motor domain head back into the forward direction for the next step. Meanwhile, the beta-linker region acts as the second part of the transducer by extending and contracting to produce the forward force. In the case of the kinesin-like proteins which are deemed non-processive, they would still be able to carry out the latter steps, and thus still carry out motility; however, their inability to release the coiled-coil domain efficiently would not result in the full extension needed to reach both sites transiently, and thus are more likely to be dislodged from the microtubule and dissociate. Moreover, the observation that the Ncd motor domain can be switched to move in the same direction as kinesin by the presence of the adjacent neck region (Case, *et al.*, 1997; Henningsen and Schliwa, 1997), but is still not processive (Case, *et al.*, 1997), indicates that, although positioning of the detached motor domain has switched, the necessary change in interactions occuring between the two motor domain heads and the neck region, which are critical for driving the proposed separation, are not present. Thus, the ability to carry out highly processive movement involves a set of highly coupled changing structural interactions between the motor domains, the beta-linker regions and the coiled-coil domain.

D. Conclusion

In summary, using several synthetic peptides corresponding to the kinesin neck region, we have clearly demonstrated that the neck coiled-coil dimerization domain exist from residues 337 to 369. Additionally, we have shown that the stability of the coiled-coil region is critically dependent upon both the first and last heptads, and can be further increased by the two N-terminal helix capping residues Leu 335 and Thr 336 within the "beta-linker" region. The latter observations, in conjunction with the conservation of the "destabilizing residues" within the hydrophobic core of the coiled-coil domain, have thus led us to propose a model for processivity in which a conformational change within one of the kinesin motor domains, driven by a change in nucleotide binding state, can be transmitted to the helix capping residues to cause a change in their orientation and/or packing thus conceivably driving the momentary unfolding of the coiled-coil domain required for separation and elongation to the next microtubule binding site. Moreover, the observation that these features largely exist only within the conventional kinesin sequences suggest that these features represent a specialized adaption that is likely critical for kinesin's highly processive nature. Overall, these results further support the premise that the neck region is an important structural element involved in processivity, directionality and motor speed.

E. References

Berger, B., Wilson, D. B., Wolf, E., Tonchev, T., Milla, M. and Kim, P. S. (1995). Predicting coiled coils by use of pairwise residue correlations. *Proc Natl Acad Sci U S A* **92**, 8259-8263.

Block, S. M. (1998). Kinesin: what gives? *Cell* **93**, 5-8.

Case, R. B., Pierce, D. W., Hom-Booher, N., Hart, C. L. and Vale, R. D. (1997). The directional preference of kinesin motors is specified by an element outside of the motor catalytic domain. *Cell* **90**, 959-966.

Correia, J. J., Gilbert, S. P., Moyer, M. L. and Johnson, K. A. (1995). Sedimentation studies on the kinesin motor domain constructs K401, K366, and K341. *Biochemistry* **34**, 4898-907.

Crevel, I. M., Lockhart, A. and Cross, R. A. (1996). Weak and strong states of kinesin and ncd. *J Mol Biol* **257**, 66-76.

Cross, R. A. (1995). On the hand-over-hand footsteps of kinesin heads. *J Muscle Res Cell Motil* **16**, 91-94.

Fischetti, V. A., Landau, G. M., Schmidt, J. P. and Sellers, P. (1993). Identifying periodic occurences of a template with applications to protein structure. *Information Processing Letters* **45**, 11-18.

Hackney, D. D. (1994). Evidence for alternating head catalysis by kinesin during microtubule- stimulated ATP hydrolysis. *Proc Natl Acad Sci U S A* **91**, 6865-6869.

Hackney, D. D. (1994). The rate-limiting step in microtubule-stimulated ATP hydrolysis by dimeric kinesin head domains occurs while bound to the microtubule. *J Biol Chem* **269**, 16508-16511.

Hancock, W. O. and Howard, J. (1998). Processivity of the motor protein kinesin requires two heads. *J Cell Biol* **140**, 1395-1405.

Henningsen, U. and Schliwa, M. (1997). Reversal in the direction of movement of a molecular motor [see comments]. *Nature* **389**, 93-96.

Hirose, K., Lockhart, A., Cross, R. A. and Amos, L. A. (1995). Nucleotide-dependent angular change in kinesin motor domain bound to tubulin [see comments]. *Nature* **376**, 277-279.

Hirose, K., Lockhart, A., Cross, R. A. and Amos, L. A. (1996). Three-dimensional cryoelectron microscopy of dimeric kinesin and ncd motor domains on microtubules. *Proc Natl Acad Sci U S A* **93**, 9539-9544.

Huang, T. G., Suhan, J. and Hackney, D. D. (1994). Drosophila kinesin motor domain extending to amino acid position 392 is dimeric when expressed in Escherichia coli. *J Biol Chem* **269**, 32708.

Jiang, W., Stock, M. F., Li, X. and Hackney, D. D. (1997). Influence of the kinesin neck domain on dimerization and ATPase kinetics. *J Biol Chem* **272**, 7626-7632.

Kammerer, R. A., Schulthess, T., Landwehr, R., Lustig, A., Engel, J., Aebi, U. and Steinmetz, M. O. (1998). An autonomous folding unit mediates the assembly of two-stranded coiled coils. *Proc Natl Acad Sci U S A* **95**, 13419-13424.

Kozielski, F., Sack, S., Marx, A., Thormahlen, M., Schonbrunn, E., Biou, V., Thompson, A., Mandelkow, E. M. and Mandelkow, E. (1997). The crystal structure of dimeric kinesin and implications for microtubule-dependent motility. *Cell* **91**, 985-994.

Lupas, A., Van Dyke, M. and Stock, J. (1991). Predicting coiled coils from protein sequences. *Science* **252**, 1162-1164.

Ma, Y. Z. and Taylor, E. W. (1997). Interacting head mechanism of microtubule-kinesin ATPase. *J Biol Chem* **272**, 724-730.

Monera, O. D., Sereda, T. J., Zhou, N. E., Kay, C. M. and Hodges, R. S. (1995). Relationship of sidechain hydrophobicity and alpha-helical propensity on the stability of the single-stranded amphipathic alpha-helix. *J Pept Sci* **1**, 319-329.

Morii, H., Takenawa, T., Arisaka, F. and Shimizu, T. (1997). Identification of kinesin neck region as a stable alpha-helical coiled coil and its thermodynamic characterization. *Biochemistry* **36**, 1933-1942.

Moyer, M. L., Gilbert, S. P. and Johnson, K. A. (1996). Purification and characterization of two monomeric kinesin constructs. *Biochemistry* **35**, 6321-6329.

Munoz, V., Blanco, F. J. and Serrano, L. (1995). The hydrophobic-staple motif and a role for loop-residues in alpha- helix stability and protein folding. *Nat Struct Biol* **2**, 380-385.

Richardson, J. S. and Richardson, D. C. (1988). Amino acid preferences for specific locations at the ends of alpha helices [published erratum appears in Science 1988 Dec 23;242(4886):1624]. *Science* **240**, 1648-1652.

Romberg, L., Pierce, D. W. and Vale, R. D. (1998). Role of the kinesin neck region in processive microtubule-based motility. *J Cell Biol* **140**, 1407-1416.

Sack, S., Muller, J., Marx, A., Thormahlen, M., Mandelkow, E. M., Brady, S. T. and Mandelkow, E. (1997). X-ray structure of motor and neck domains from rat brain kinesin. *Biochemistry* **36**, 16155-16165.

Seale, J. W., Srinivasan, R. and Rose, G. D. (1994). Sequence determinants of the capping box, a stabilizing motif at the N- termini of alpha-helices. *Protein Sci* **3**, 1741-175.

Steinmetz, M. O., Stock, A., Schulthess, T., Landwehr, R., Lustig, A., Faix, J., Gerisch, G., Aebi, U. and Kammerer, R. A. (1998). A distinct 14 residue site triggers coiled-coil formation in cortexillin I. *Embo J* **17**, 1883-1891.

Thormahlen, M., Marx, A., Sack, S. and Mandelkow, E. (1998). The coiled-coil helix in the neck of kinesin. *J Struct Biol* **122**, 30-41.

Tripet, B., Chana, M. and Hodges, R. S. (1999). STABLECOIL: A coiled-coil prediction method based on stability co-efficients. *J. Biol Chem.* (Summited).

Tripet, B., Vale, R. D. and Hodges, R. S. (1997). Demonstration of coiled-coil interactions within the kinesin neck region using synthetic peptides. Implications for motor activity. *J Biol Chem* **272**, 8946-8956.

Chapter VI

De novo design of a model peptide sequence to examine the effects of single amino acid substitutions in the hydrophobic core on both stability and oligomerization state of coiled-coils

A version of this chapter has been published: Kurt Wagschal, Brian Tripet and Robert S. Hodges, *J. Mol. Biol.* (1999) 285, 785-803.

A. Introduction

The coiled-coil structural motif, first proposed by Crick in 1953 based on the x-ray diffraction data of α-keratin (Crick, 1953), consists of right-handed amphipathic α-helices that adopt a left-handed supercoil, such that their hydrophobic faces are in continuous contact along the length of the coiled-coil. Analysis of the sequence of the two-stranded coiled-coil tropomyosin revealed that each α-helix consists of a regularly repeating heptad unit designated *abcdefg*, in which positions "**a**" and "**d**" are occupied by predominately hydrophobic amino acids in the hydrophobic core of the protein fold (Hodges *et al.*, 1972; Sodek *et al.*, 1972). The predicted structure for a three-stranded coiled-coil was confirmed in 1981 by the high resolution x-ray structure of influenza virus hemagglutinin (Wilson *et al.*, 1981), while the first high resolution x-ray structure of a two-stranded coiled-coil representing the dimerization site of the DNA-binding protein GCN-4 was reported in 1991 (O'Shea *et al.*, 1991). Since then over 20 other proteins containing coiled-coil domains have been solved to high resolution (Lupas, 1996). To date over 200 proteins have been identified as containing the coiled-coil oligomerization domain, representing a wide range of applications in nature for this important structural motif (recent reviews see Adamson *et al.*, 1993; Alber, 1992; Cohen & Parry, 1990; Hodges, 1996; Kammerer, 1997; Kohn & Hodges, 1998).

It has been firmly established that the coiled-coil hydrophobic core "**a**" and "**d**" positions are in large part responsible for defining the stability and oligomerization state of the structure (Betz *et al.*, 1995; Cedervall *et al.*, 1997; Greenfield & Hitchcock-DeGregori, 1995; Harbury *et al.*, 1993; Hodges, 1992; Monera *et al.*, 1996; Zhou *et al.*, 1992a,b; Zhu *et al.*, 1993), and that substitution of polar, hydrophilic residues in these positions sometimes results in unique oligomerization states at the expense of overall stability (Gonzales *et al.*, 1996a,b; O'Shea *et al.*, 1991). It follows then that both the *de novo* design of a coiled-coil sequence with a specific stability and oligomerization state, and the identification and preliminary characterization of a coiled-coil structural motif in a naturally occurring protein sequence require knowledge of the contribution each amino acid residue can make toward these properties when present at an "**a**" or "**d**" position. Presently, however, no systematic and comprehensive study has yet been carried out using well-defined coiled-coils to determine the contribution of each side chain in the hydrophobic core towards stability and oligomerization state. Thus, we describe here the *de novo* design and complete biophysical characterization of a model coiled-coil sequence, termed X19**a** (Figure VI-1), as a first step toward determining these properties for position "**a**" analogs. Peptide sequence X19**a** analogs I19**a**, L19**a**, N19**a**, A19**a**, E19**a**, and C36**d**L19**a** (Figure VI-1) were strategically selected to delineate the ability of peptide sequence X19**a** to measure the entire spectrum of stabilities that would be expected to arise in an analog set comprising all 20 naturally occurring amino acids. We also describe the systematic biophysical characterization necessary to ensure that each analog underwent chemical denaturation in a concentration-independent manner from a parallel two-stranded disulfide-bridged α-helical coiled-coil (monomer) regardless of the oligomerization state in benign medium (buffer in the absence of denaturant). Thus this allows the stability of analogs having different oligomerization states in benign buffer to be directly compared since unfolding occurs from the same monomeric state. Hence, the stability data obtained in this study is only relevant to two-stranded coiled-coils. In general, this is the first study to provide a model system for systematically examining the effect of amino acid side-chain substitution in the hydrophobic core of a folded protein on both stability and oligomerization state.

B. Results and Discussion

Peptide design

Figure VI-1 shows the amino acid sequences and nomenclature of the synthetic peptides that were used in this study. These analogs were analyzed with the N-terminal Cys side-chain thiol functional group reduced, denoted X19**a**(red), with the Cys thiol functionality derivatized to form the *S*-carboxamidomethyl cysteine group and thereby preclude spurious oxidation, denoted X19**a**(cm), and also with the Cys residues oxidized to form an interchain disulfide bond, denoted X19**a**(ox). Coiled-coil sequence X19**a** was designed to serve as a scaffold in which to carry out amino acid substitution at hydrophobic core position 19**a** in the central heptad to measure relative hydrophobicity, thermodynamic stability and oligomerization state. Substitutions were made in the central heptad since perturbing the hydrophobic core in this region has been shown to have the greatest effect on coiled-coil stability (Zhou *et al.*, 1992b). Moreover, the crucial role of the central heptad in controlling oligomerization state has been demonstrated by O'Shea *et al.* (1991), where they showed that a central Asn residue of the transcription activator GCN-4 directs the molecule to a monomeric, two-stranded coiled-coil. In general, the information obtained from these analogs will be applicable to physiologically important coiled-coil structures since peptide sequence X19**a** is similar to bZip transcription factors in length (35-38 residues) and contains a hydrophobic core with Leu at the "**d**" positions and the β-branched amino acid Val at the "**a**" positions (Alber, 1992; Hu & Sauer, 1992).

The quatenary structures of peptide sequence X19**a** oxidized, reduced or carboxamidomethylated monomer, dimer and trimer oligomerization states are shown in

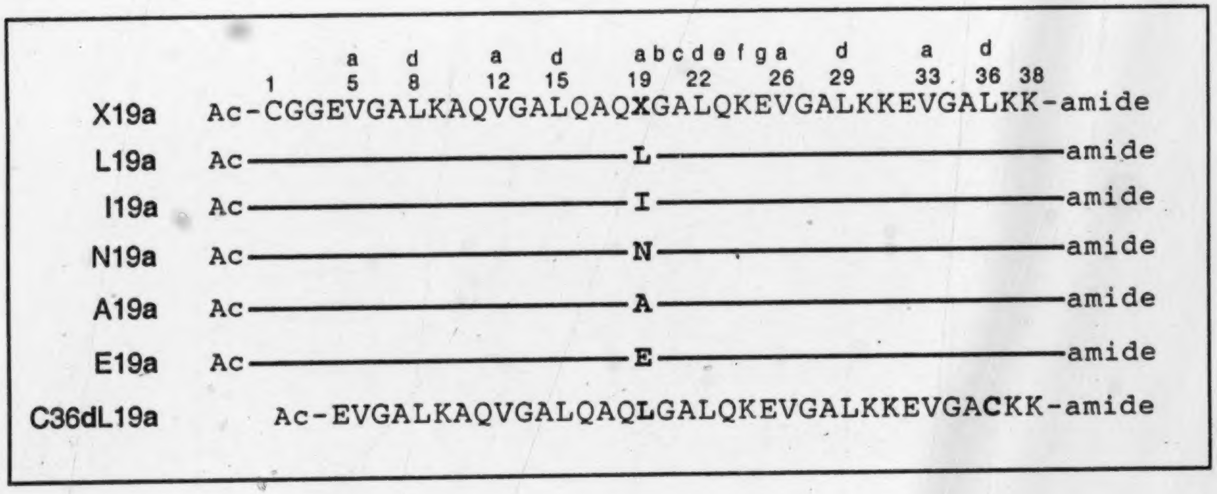

a d a d abcdefga d a d
1 5 8 12 15 19 22 26 29 33 36 38
X19a Ac-CGGEVGALKAQVGALQAQXGALQKEVGALKKEVGALKK-amide
L19a Ac L amide
I19a Ac I amide
N19a Ac N amide
A19a Ac A amide
E19a Ac E amide
C36dL19a Ac-EVGALKAQVGALQAQLGALQKEVGALKKEVGACKK-amide

Figure VI-1. The amino acid sequence and nomenclature of the Leu, Ile, Asn, Ala and Glu model X19a analogs examined in this study. Sequence C36dL19a is an analog of L19a in which the position of the disulfide bond was changed from the N-terminus of the peptide sequence to the C-terminal d position (position 36d).

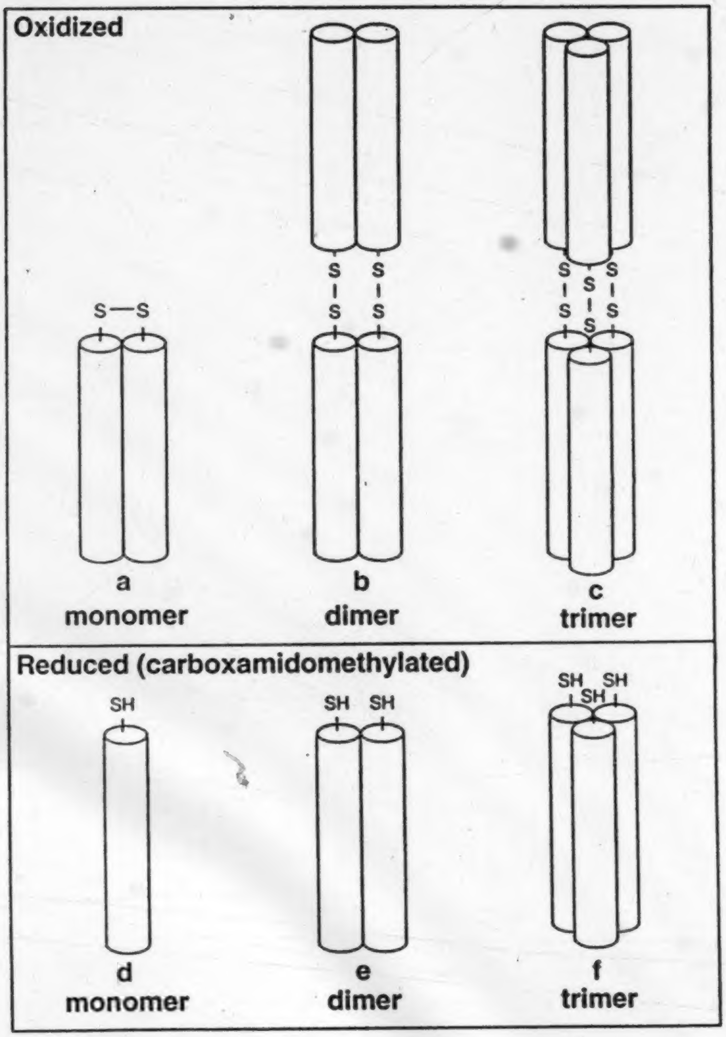

Oxidized
S—S
S
S
S
S
S
S
S
S
S
S
a
monomer
b
dimer
c
trimer
Reduced (carboxamidomethylated)
SH
SH SH
SH SH
SH
d
monomer
e
dimer
f
trimer

Figure VI-2. Top panel, proposed models of the interchain disulfide-bridged peptides (oxidized): monomer (a), dimer (b), and trimer (c).The dimer of oxidized peptides was not observed here. Bottom panel, proposed models of the peptides when the disulfide bridge is reduced or carboxamidomethylated: monomer (d), dimer (e), and trimer (f). The cylinders represent the 35-residue polypeptide chain folded as an α-helix.

Figure VI-2. Thus, the oxidized monomer consists of a single two-stranded coiled-coil molecule, where the two α-helices are disulfide-bridged at the amino terminus via a flexible Cys-Gly-Gly linker, which serves to orient the α-helices in a parallel manner (Figure VI-2a). The reduced monomer (Figure VI-2d) consists of a single polypeptide chain, which can fold as a single α-helix on a hydrophobic matrix during reversed-phase HPLC. The oxidized dimer (Figure VI-2b) consists of two molecules, where each molecule contains two α-helical polypeptide chains linked by a disulfide bridge, forming a two-stranded coiled-coil. The reduced dimer (Figure 2e) consists of two α-helices associated in a parallel manner, due to favorable electrostatic interactions, to form a two-stranded coiled-coil. The oxidized trimer is an elongated three-stranded structure, consisting of three molecules, where each molecule contains two α-helical polypeptide chains linked by a disulfide bridge (Figure VI-2c). The reduced trimer consists of three molecules, where each molecule contains one α-helical polypeptide chain, which are oriented in a parallel manner due to favorable electrostatic interactions (Figure VI-2f).

Figure VI-3 shows the helical wheel diagram of the proposed three-stranded coiled-coil formed by X19**a**, while Figure VI-4 panel A shows the helical wheel diagram of the analogous two-stranded coiled-coil. The two-headed arrows in the hydrophobic core depict Van der Waals interactions between "**a**" and "**d**" residues on adjacent chains. The two-headed dashed arrows show the location of electrostatic interactions formed between Lys and Glu residues at positions "**e**" and "**g**" of adjacent α-helices which, in addition to providing ~ 0.5 kcal/mol towards stability (Zhou *et al.*, 1994), serve to orient the polypeptide chains in an in-register and parallel manner. Three Lys residues were placed at the solvent exposed positions "**f**' of each α-helix to increase the solubility of the peptides at pH 7.0 and to discourage non-specific aggregation. The uncharged residues Gly and Ala were placed at positions "**b**" and "**c**" to avoid intrahelical ionic interactions with adjacent residues at positions "**e**" and "**g**". As seen in Figure VI-4 panel B, which shows a side view of model sequence X19**a** in a two-stranded coiled-coil, the hydrophobic

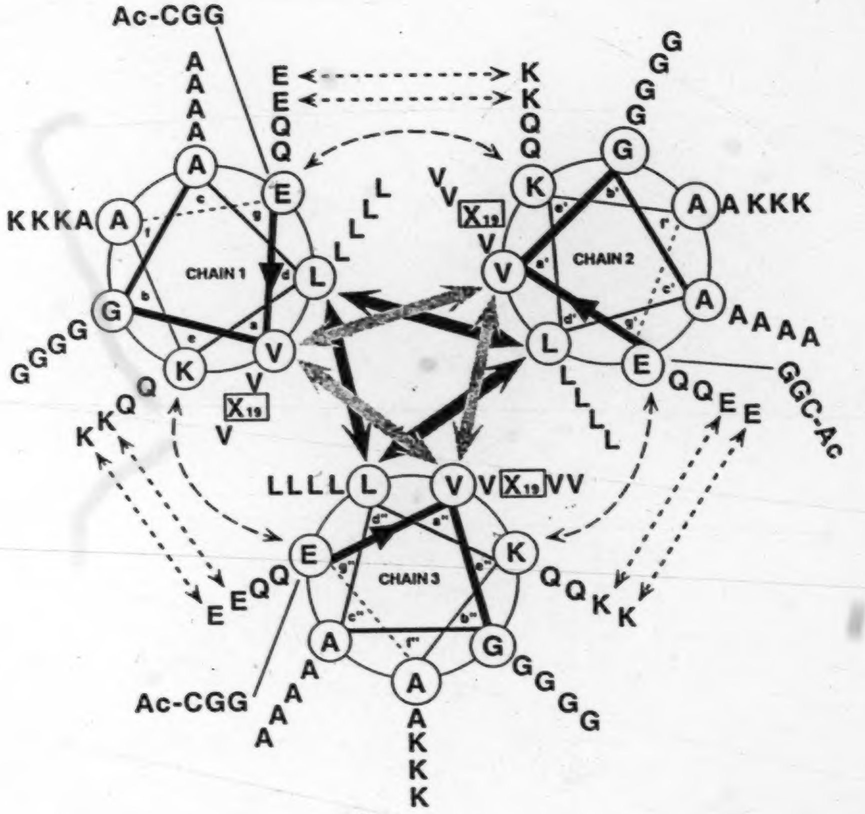

Ac-CGG
CHAIN 1
CHAIN 2
CHAIN 3
X19
KKKA
A KKK
GGGG
AAAA
EEQQ
KKQQ
QQEE
QQKK
LLLL
VVVX19VV
GGC-Ac
Ac-CGG
AKKK
a
b
c
d
e
f
g
a'
b'
c'
d'
e'
f'
g'
a"
b"
c"
d"
e"
f"
g"

Figure VI-3. End view of proposed X19a three-stranded reduced coiled-coil from the N-terminus. The direction of propagation of the peptide chains is into the page from the N-terminus to the C-terminus, with the chains parallel and in-register. Residues in the first two helical turns are circled. Heptad positions are labelled a-g, with the prime and double-prime indicating corresponding positions on the second and third α-helices of the trimer. The site of substitution at position 19 is boxed and denoted by an "X". Wide arrows in the core depict the hydrophobic interaction that occurs between residues in the a and d positions. Interhelical electrostatic interactions between residues in e and g positions are indicated as dashed arrows.

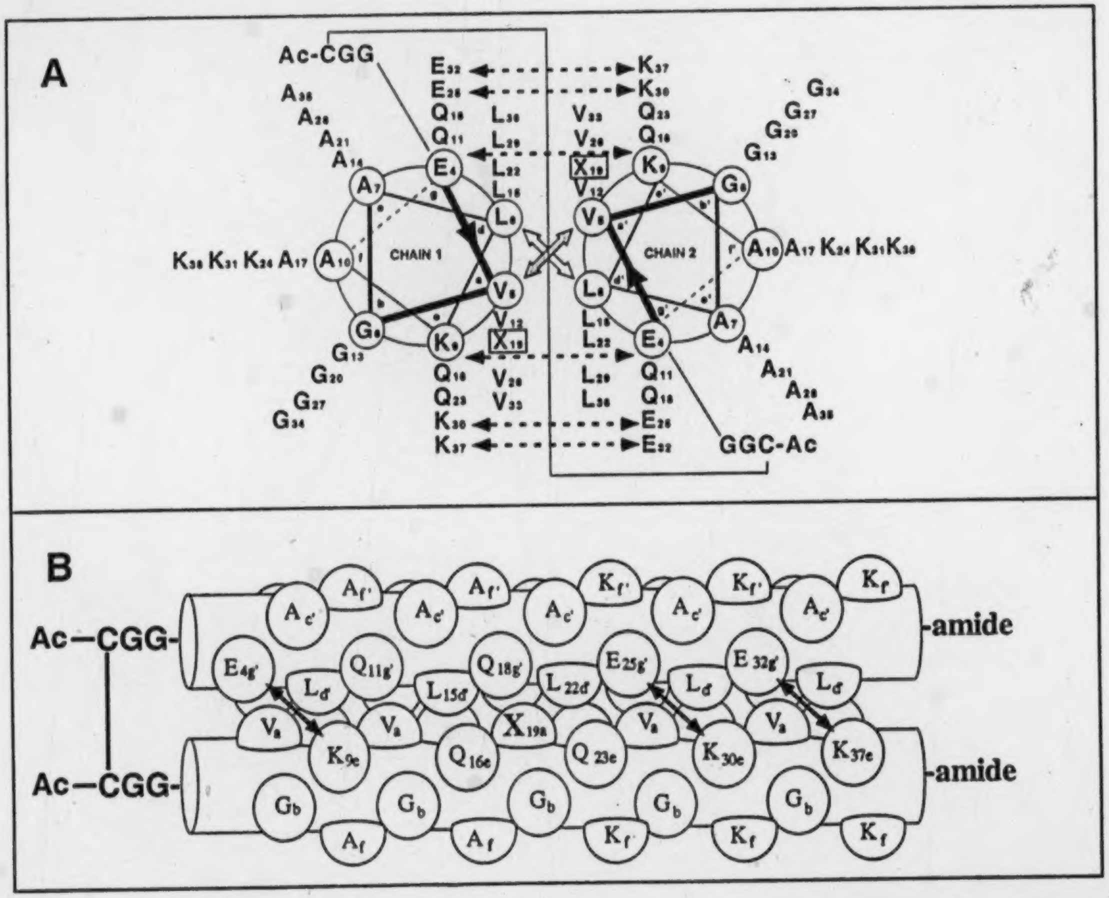

A
Ac-CGG
GGC-Ac
CHAIN 1
CHAIN 2
E4
A7
A10
G6
K9
V8
L4
X19
V12
K38 K31 K24 A17
A10 A17 K24 K31 K38
A35 A28 A21 A14
G13 G20 G27 G34
E32 E25 Q18 Q11
K37 K30 Q23 Q16
L36 L29 L22 L15
V33 V26
B
Ac—CGG-
-amide
Ac'
Af'
Kf'
Kf
E4g'
Q11g'
Q18g'
E25g'
E32g'
Ld'
L15d'
L22d'
Va
X19a
K9e
Q16e
Q23e
K30e
K37e
Gb
Af

Figure VI-4. End and side views of the structure proposed for X19a in an oxidized, two-stranded monomeric oligomerization state. Panel A, view is from the N-terminus of the oxidized monomer, and the nomenclature is the same as in Figure 2. Panel B, side view of the oxidized monomer. Amino acids defining the micro-environment surrounding the site of amino acid substitution at position 19a (X19a) are shaded. Interchain electrostatic interactions occurring across the hydrophobic face are indicated by double-headed arrows.

environment around the site of substitution on this side of the coiled-coil is largely defined by Leu15**d'** and Leu22**d'** (shaded) and by Gln residues strategically placed at positions 16**e**, 18**g'** and 23**e** (shaded). Gln residues were chosen for these positions to provide moderately polar, but uncharged, side-chains that prevent charge-charge interactions with the substituted amino acid side-chain at position 19**a**. These Gln residues along with Gln11**g'** also preclude the formation of interhelical **e-g'** salt bridges around the substitution site. In addition, the N-termini of the peptides were acetylated and the C-termini amidated in order to eliminate the introduction of charges and possible unfavourable interactions with the α-helix dipole (Shoemaker *et al.*, 1987).

Since two- and three-stranded coiled-coils represent the most common oligomerization states of naturally occurring coiled-coils (Lupas, 1996), the hydrophobic core residues were chosen to form two- and three-stranded structures that allow modulation of the oligomerization state by amino acid substitution at position 19**a**. DeGrado and coworkers have shown previously by x-ray crystallography that their model coiled-coil, termed Coil V_a-L_d which contains Val in each **"a"** position and Leu in each **"d"** position (Betz *et al.*, 1995), forms an all-parallel trimeric structure (Ogihara *et al.*, 1997). Also, Harbury *et al* (1993) have shown that a coiled-coil with the same hydrophobic core as Coil V_a-L_d formed a two-stranded coiled-coil when disulfide-bridged, and both two-and three-stranded coiled-coils in the absence of a disulfide bridge. Moreover, analysis of various *de novo* designed peptides utilizing a similar hydrophobic core in our own laboratory adopt three-stranded coiled-coils (Kohn *et al.*, 1998). Based on these results the hydrophobic core of model X19**a** contained Val in each **"a"** position and Leu in each **"d"** position.

In order to construct a relative stability scale for a complete set of analogs it was necessary that denaturation occur via a two-state process from folded disulfide-bridged monomer (m_f) to the corresponding unfolded monomer (m_u) regardless of the oligomerization state in benign medium. Thus, a flexible Cys-Gly-Gly linker was therefore added to the N-termini to permit disulfide bond formation and unconstrained helical alignment, since it had been shown previously that a disulfide bridge in the hydrophobic core itself can disrupt or modulate interhelical packing depending on location(Zhou *et al.*, 1993). This also has the advantage that the protein stability measurements are then concentration independent, and therefore more precise, and also ensures a parallel, in-register coiled-coil for the monomeric species. The ability of the model sequence to form a trimeric structure in benign medium, and yet denature from a monomeric state in the event a trimeric structure is indeed formed, are reconciled if the oxidized coiled-coil trimer undergoes a transition to an oxidized coiled-coil monomer at low concentrations of denaturant prior to a change in α-helical content. This has in fact been observed previously in our laboratory for an oligomeric *de novo* designed coiled-coil (Zhu *et al.*, 1993). It has previously been shown that the position of the disulfide bond can influence the resulting oligomerization state (Zhou *et al.*, 1993). To test this mode of oligomerization state control in the present study, peptide sequence C36**d**L19**a** was designed in which the position of the disulfide bond was changed by substituting Cys at residue 36 in the C-terminal hydrophobic core "**d**" position.

Choice of analogs

The X19**a** analogs L19**a**, I19**a**, N19**a**, A19**a**, and E19**a** were chosen for synthesis to evaluate the utility of the sequence for obtaining both a relative hydrophobicity scale and a relative stability scale for an entire set of analogs resulting from substitution of all naturally-occurring amino acids, as well as for investigating the effect of single amino acid substitution on oligomerization state. These analogs allowed a thorough assessment of the sequence since the overall stability must be high enough such that disruptive, charged side-chains such as that of Glu can be incorporated and still result in formation of a coiled-coil structure. Conversely, substitution of highly hydrophobic residues such as Leu and Ile will not result in a coiled-coil structure that is too stable to allow determination of transition midpoints during chemical denaturation. Also, besides these limits of stability that must be adhered to, the hydrophobicity and stability must be modulated over a large enough range to result in an accurate relative scale. A19**a** was examined since it has an intermediate hydrophobicity (Monera *et al.*, 1995), making it an ideal reference for calculating relative hydrophobicity and stability. Extensive studies have been published which show that an Asn residue at the "**a**" position in the central heptad of a coiled-coil will lead exclusively to a two-stranded structure (O'Shea *et al.*, 1991), and N19**a** was synthesized to test whether Asn has the same effect in peptide sequence X19**a**. Synthesis of both L19**a** and the β-branched side-chain isostere I19**a** allowed determining not only the hydrophobicity and facility of denaturation for the most hydrophobic analogs, but also assessment of subtle packing effects on overall stability and oligomerization state since it has been previously shown that β-branched amino acids pack more efficiently in a hydrophobic core "**a**" position than in a hydrophobic core "**d**" position (Zhu *et al.*, 1993).

CD spectroscopy

The CD spectra of each oxidized analog was measured at 25 °C under physiologically relevant, benign conditions (100 mM KCl, 50 mM PO_4, pH 7.0), and also in 50% TFE for all analogs (Table VI-1). The estimated maximal theoretical molar ellipticity at 222 nm and 25 °C is -38,200 for the 76-residue peptide X19**a**(ox) (Chen *et al.*, 1974), and the analogs generally approached this maximal value under benign conditions at pH 7.0. Under benign conditions the ratio of ellipticities 222 nm/208 nm was greater than 1.0 for all peptides, which is indicative of the presence of interacting α-helices (Cooper & Woody, 1990; Lau *et al.*, 1984). It was observed that the molar ellipticity in 50% TFE at 222 nm generally decreased by about 10%. This decrease in helicity in TFE may reflect the loss of synergism in the n-π^* transition (222 nm CD band) from interacting α-helices in the coiled-coil structure (Lau *et al.*, 1984) since TFE has

Table VI-1. CD spectroscopy results for coiled-coil X19a analogs

Peptide^a	$[\Theta]$ 222^b **Benign** (degrees•cm^2•dmol^{-1})	% α-Helix^c	$[\Theta]$ 222/208 ^{d} **Benign**	$[\Theta]$ 222 ^{e} **50%TFE** (degrees•cm^2•dmol^{-1})	$[\Theta]$ 222/208 ^{d} **50%TFE**
L19a(ox)	-36600	105%	1.05	-33300	0.89
I19a(ox)	-36100	104%	1.06	-30800	0.90
A19a(ox)	-34800	100%	1.08	-30200	0.85
N19a(ox)	-34500	99%	1.03	-33600	0.91
E19a(ox)	-35900	103%	1.07	-25400	0.83
C36dL19a(ox)	-33700	97%	1.03	-29000	0.87

[a] Name of peptides. See Figure VI-1 for peptide nomenclature and sequences.

[b] Mean residue molar ellipticities at 222 nm. Measurements were performed on ~50 μM samples of oxidized (disulfide-bridged) peptide in benign buffer (100 mM KCl, 50 mM PO_4, pH 7.0) at 25 °C.

[c] The percent α-helical content was calculated based on a predicted molar ellipticity of -34,730 at 25 °C for a 35-residue helical peptide (Chen *et al.*, 1974; Chang *et al.*, 1978).

[d] A ratio of $[\theta]$222/208 greater than 1.0 is a measure of interacting α-helices in benign media, and a ratio of ~0.90 represents non-interacting α-helices in 50% TFE (Lau *et al.*, 1984).

[e] Mean residue molar ellipticities at 222 nm. Measurements were perfomed on ~50 μM samples of oxidized peptide in benign buffer diluted 1:1 (v/v)with TFE.

Table VI-2. Biophysical data for the model X19a coiled-coil analogs.

Peptide[a]	Δt_R(Ala)[b] (min)	$[\mathrm{GdnHCl}]_{1/2}[\Theta]_{222}$[c] (M)	$[\mathrm{GdnHCl}]_{1/2}$HPSEC[d] (M)	Slope[e] *m* (kcal/mol•M)	$\Delta\Delta G_{(\mathrm{Ala})}$[f] (kcal/mol)	Oligomer ratio[g]
L19a(ox)	+7.9	2.9	2.7	1.68	-3.4	t
I19a(ox)	+7.6	3.3	3.2	1.31	-3.7	3:2 (t:m)
A19a(ox)	0	1.2	1.0	2.20	0	2:3 (t:m)
N19a(ox)	-6.8	1.5	1.6	1.84	-0.77	m
E19a(ox)	-8.8	0.52	n.d.[h]	2.70	+1.6	7:3 (t:m)

[a] Name of peptides. See Figure VI-1 for peptide nomenclature and sequence.

[b] Difference in the RP-HPLC retention time of X19a relative to the Ala substituted analog (A19a), where the cysteine residue in each peptide was derivatized to *S*-carboxamidomethyl cysteine.

[c] $[\mathrm{GdnHCl}]_{1/2}[\Theta]_{222}$ is the transition midpoint obtained from GdnHCl denaturation experiments monitored by CD using molar ellipticity at 222 nm, and is the concentration of GdnHCl (M) required to give a 50% decrease in molar ellipticity. Performed at 50μM peptide concentrations.

[d] $[\mathrm{GdnHCl}]_{1/2}$HPSEC is the transition midpoint obtained from GdnHCl denaturation experiments monitored by HPSEC, and is the concentration of GdnHCl required to shift the retention time of HPSEC peak m_x midway between the HPSEC retention times of m_f and m_u (figures 9 and 10). Peak m_x results from rapid equilibration of folded monomer m_f and unfolded monomer m_u, and its retention time is dependent on [GdnHCl]. Performed at 2.5 μM peptide concentrations.

[e] *m* is the slope term in the equation $\Delta G = \Delta G_u^{H2O} + m$[denaturant] (Pace, 1986), used in the calculation of $\Delta\Delta G_{(\mathrm{Ala})}$.

[f] $\Delta\Delta G_{(\mathrm{Ala})}$ is the difference in free energy of unfolding of each analog relative to the Ala substituted analog, and was calculated from the equation: $\Delta\Delta G_{(\mathrm{Ala})} = \{(m_{\mathrm{Ala}} + m_x)/2\}\{([\mathrm{denaturant}]_{1/2})_x - ([\mathrm{denaturant}]_{1/2})_{\mathrm{Ala}}\}$ (Sali *et al.*, 1991).

[g] Ratio of oligomerization states present by analyzing an 8 μL aliquot of a 500 μM peptide solution by HPSEC as described in the methods. m denotes a monomeric oligomerization state, and t denotes a trimeric oligomerization state.

[h] n.d., not determined.

been shown to disrupt tertiary and quartenary structure, and to promote secondary structure (Cooper & Woody, 1990; Lau *et al.*, 1984; Sönnichsen *et al.*, 1992). Interestingly, the negatively charged Glu residue was much more disruptive of α-helical structure as a monomeric α-helix in the non-polar environment of TFE than in the coiled-coil structure in benign medium (Table VI-1). The ratio of ellipticities 222 nm/208 nm in 50% TFE decreased to ~0.9 for all of the analogs, which is characteristic of non-interacting α-helices. Thus, these data taken together show that these peptides assumed coiled-coil structures under benign conditions.

Side-chain hydrophobicity

Since it has been shown that the hydrophobicity of the "**a**" and "**d**" positions provides the greatest contribution to overall stability (Cedervall *et al.*, 1997; Greenfield & Hitchcock-DeGregori, 1995; Zhu *et al.*, 1993; Zhou *et al.*, 1992a), the utility of model sequence X19**a** to measure the relative side-chain hydrophobicity of the analogs was assessed using C-8 reversed-phase HPLC. It was found that all of the X19**a** analogs were well resolved under the elution conditions employed, with a 17.7 min range in retention time between the least hydrophobic analog (E19**a**(cm), Δt_R(Ala) -8.8 min) and most hydrophobic analog (L19**a**(cm),Δt_R(Ala) +7.9 min) (Table VI-2). Moreover, the relative hydrophobicity of each side-chain agreed with the relative order of hydrophobicity previously observed in other model systems (Monera *et al.*, 1995), which indicates side-chain hydrophobicity is model-independent. These results show that peptide sequence X19**a** will allow measurement of the relative hydrophobicities of an analog set comprising all 20 naturally occurring amino acids.

Protein stability

To determine the relative stability of each of the analogs, each peptide was denatured by GdnHCl and the change in molar ellipticity followed at 222 nm. The

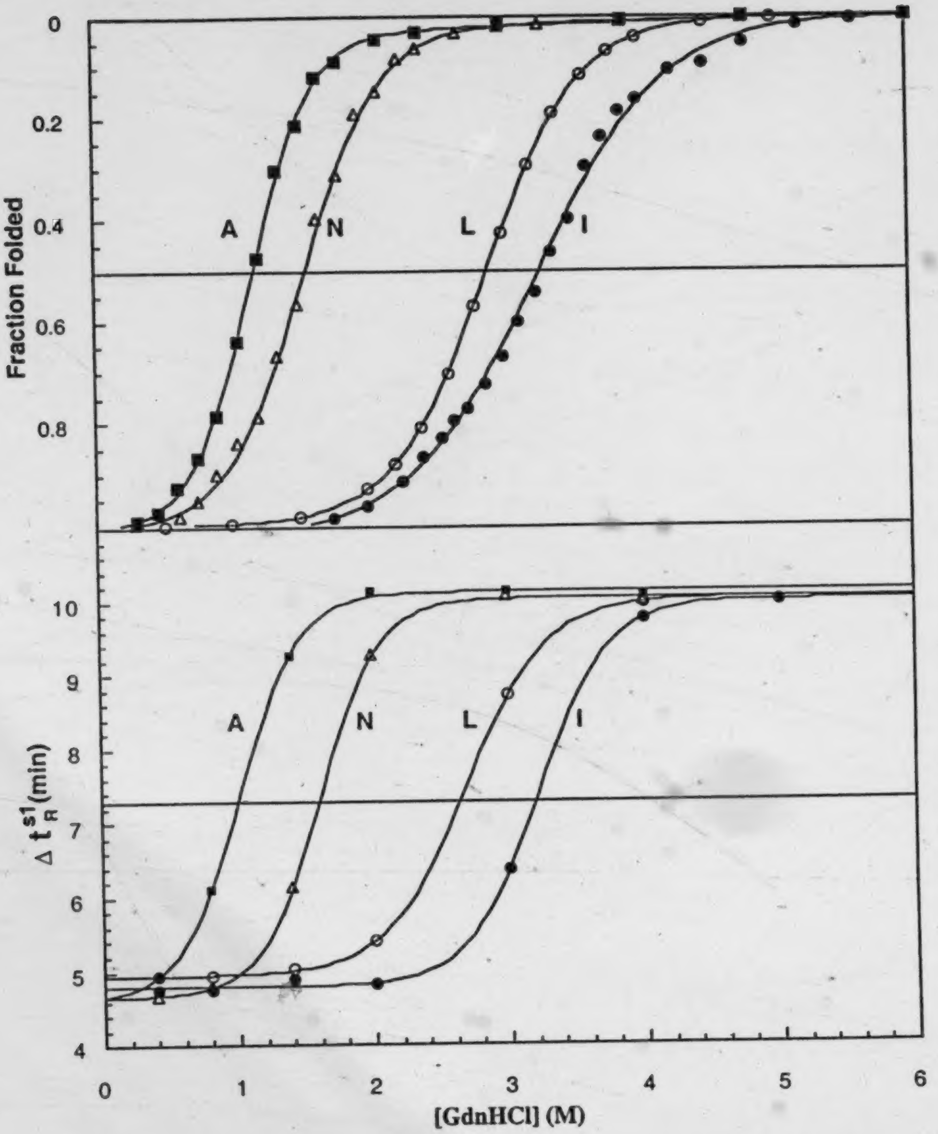


Figure VI-5. Top panel, GdnHCl denaturation profiles for the analogs monitored by change in molar ellipticity using CD. The midpoints of the GdnHCl denaturation profiles obtained by measuring a change in ellipticity at 222 nm ($[\text{GdnHCl}]_{1/2}[\theta]_{222}$) for A19a, N19a, L19a and I19a, were 1.2 M, 1.5 M, 2.9 M and 3.3 M, respectively. Bottom panel, the change in retention time of the folded monomer peak (m_f), the peak resulting from rapid exchange of the folded monomer peak and the unfolded monomer peak (m_x), and finally the unfolded monomer peak (m_u) with increasing GdnHCl concentration. The midpoints of the GdnHCl denaturation profiles obtained using this method ($[\text{GdnHCl}]_{1/2}$HPSEC) for A19a, N19a, L19a, and I19a were 1.0 M, 1.6 M, 2.7 M and 3.2 M respectively.

GdnHCl denaturation curves were all sigmoidal, with a single inflection point and cooperative transition consistent with a two-state equilibrium with no intermediate states (Figure VI-5, top panel). We observed that the $[\mathrm{GdnHCl}]_{1/2}[\theta]_{222}$ values measured for the analogs ranged from 0.52 M to 3.3 M (Table VI-2). The free energy of unfolding of the analogs relative to the Ala substituted analog, $\Delta\Delta G_{\mathrm{Ala}}$, was calculated as described in the methods section, and it was found that I19**a**(ox) was the most stable ($\Delta\Delta G_{\mathrm{Ala}} = 3.7$ kcal/mol), and E19**a**(ox) the least stable ($\Delta\Delta G_{\mathrm{Ala}} = -1.7$ kcal/mol) of the model X19**a** analogs synthesized for this study. Thus, substituting a single position "**a**" amino acid in the centre heptad of X19**a** was shown to modulate overall stability over a 5.3 kcal/mol range. This stability change correlates with the hydrophobicity change; in other words, the most hydrophobic side-chain substitution results in the most stable analog and the most hydrophilic side-chain substitution results in the least stable analog. With a view toward synthesizing an analog set comprising all naturally occurring amino acids, these results show that the model sequence will allow measurement of the stability resulting from substitution of residues imparting both high and low stability.

Oligomerization state

To assess whether the oligomerization state of the model sequence was responsive to single amino acid substitutions, each of the five analogs was analyzed by HPSEC to determine the oligomerization state (Table VI-2). Under benign buffer conditions N19**a**(ox) was monomeric and L19**a**(ox) was trimeric. The molecular weight of the HPSEC peaks assigned as monomers or trimers was corroborated by sedimentation equilibrium ultracentrifugation (SE), which is independent of molecular shape. These results confirmed that N19**a**(ox) was best fit as a single species monomer (calculated MW 7506, observed MW 7744, Figure VI-6) and that L19**a**(ox) was best fit as a single species trimer (calculated MW 22515, observed MW 22239, Figure VI-6) throughout the concentration range imposed by the sedimentation equilibrium run (~50 μM to ~500 μM).

In contrast to L19**a**(ox) and N19**a**(ox), it was observed by HPSEC that I19**a**(ox), A19**a**(ox) and E19**a**(ox) were structurally degenerate and consisted of mixtures of monomers and trimers, with monomer:trimer ratios of 3:2, 2:3, and 7:3, respectively. A major problem in *de novo* designing an amino acid sequence is limiting the molecule to fold to a single structure (Dill *et al.*, 1995). The observation that amino acid substitution in structurally unique coiled-coils can lead to structural degeneracy has previously been reported for the two-stranded GCN4-p1 leucine zipper. For example, it was found that Asn16Val, which results in a hydrophobic core with 4 Val residues in the "**a**" position and 4 Leu residues in the "**d**" position (similar to peptide sequence X19**a**), forms a mixture of two- and three-stranded species (Harbury *et al.*, 1993). Similarly, mutation of Asn16 to aminobutyric acid (Abu) also led to both two- and three-stranded oligomerization states (Gonzales *et al.*, 1996b). The fact that peptide E19**a** formed both monomeric and trimeric oligomerization states was surprising since the Glu side-chain is ionized under the pH 7.0 conditions employed, and one might expect burying three of these charged side-chains in a trimeric oligomerization state to be destabilizing. The effect of amino acid substitution on the oligomerization state of the reduced peptides was determined by examining the *S*-carboxamidomethyl cysteine derivatives L19**a**(cm), N19**a**(cm) and A19**a**(cm) by SE. These peptides are expected to behave like the corresponding reduced analogs, with the advantage that the formation of an interchain disulfide bond is prevented. The possibility that carboxamidomethylation of the cysteine residues could affect the resultant oligomerization state is minimal since the derivatized cysteine residues are not in the hydrophobic core or α-helical region due to the flexible Gly-Gly linker prior to the N-terminal cysteine. It was found that N19**a**(cm) formed mainly dimers (single species fit, calculated MW 7724, observed MW 6491, Figure VI-7), L19**a**(cm) formed exclusively trimers (single species fit, calculated MW 11583, observed MW 11293, Figure VI-7), while A19**a**(cm) formed a mixture of species ranging from dimers to trimers (data not shown). Overall, the sensitivity of the oligomerization state to the amino acid side-chain

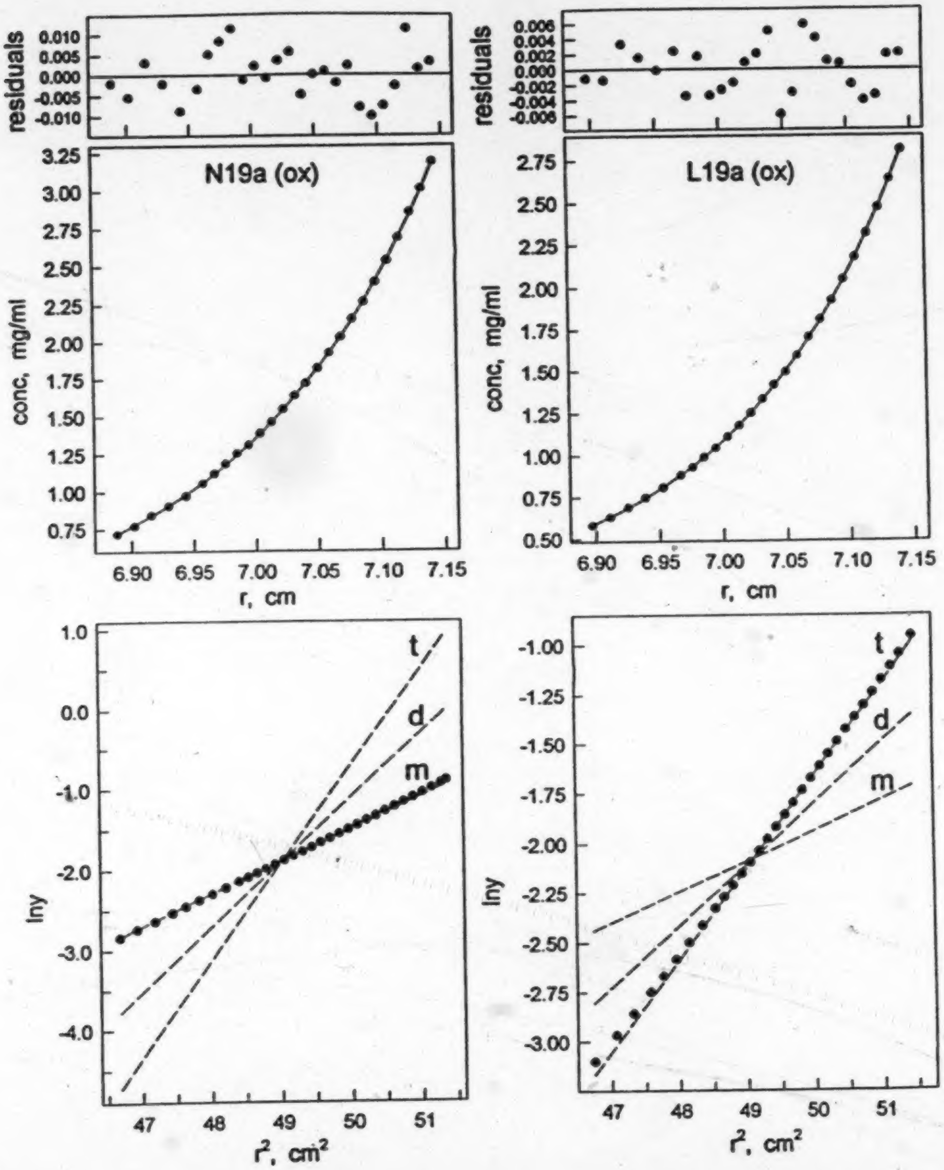


Figure VI-6. Sedimentation equilibrium analysis of N19a(ox) and L19a(ox). The middle panels (concentration versus radius) show the best fits for a single species monomeric curve fit for N19a(ox) and a single species trimeric curve fit for L19a(ox). The upper panels show the residuals from the curve fits, and the lower panels show lnY versus. r^2 of the data (filled circles) compared to theoretical monomer (m), dimer (d) and trimer (t) single species plots (brocken lines), where y is the fringe displacement, which is proportional to the concentration.

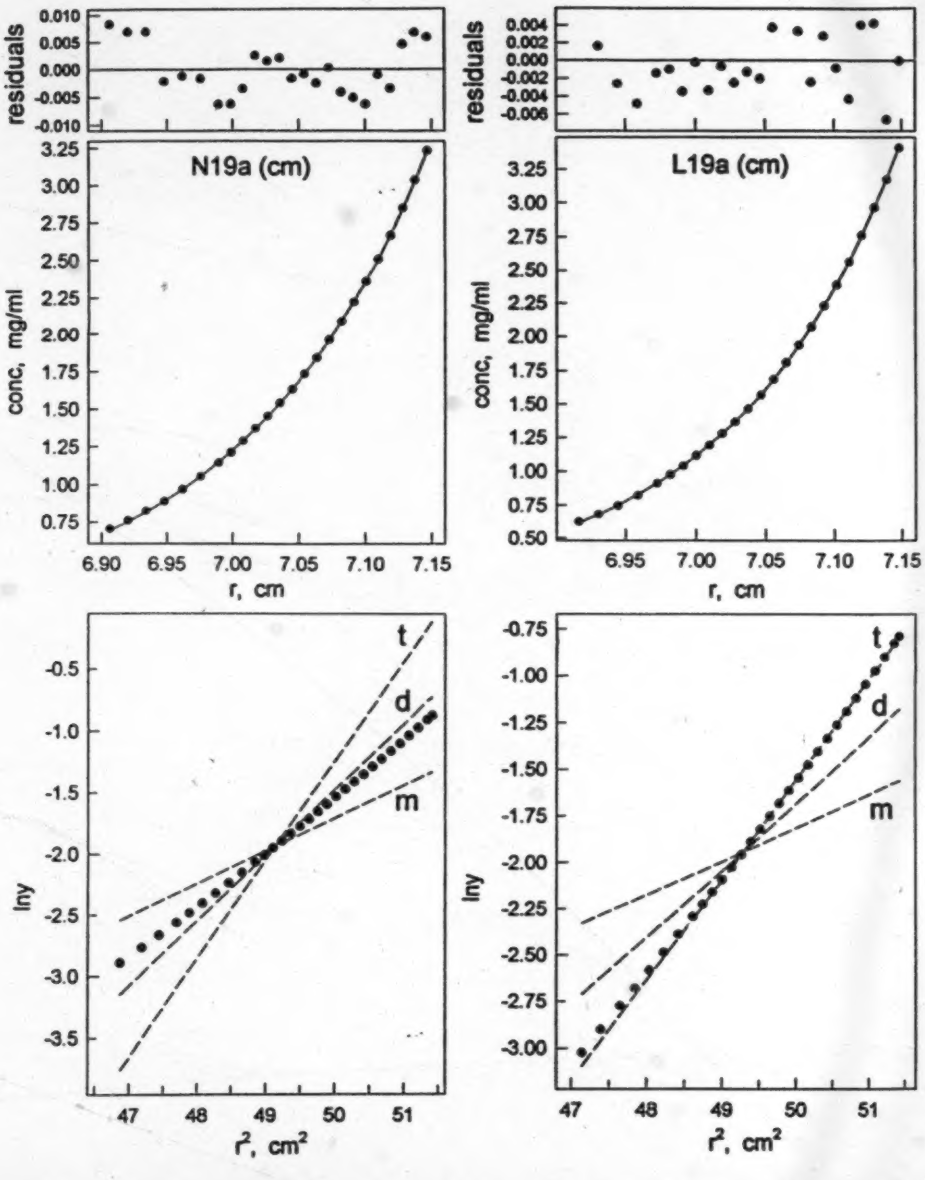


Figure VI-7. Sedimentation equilibrium analysis of N19a(cm) and L19a(cm). The middle panels (concentration versus radius) show the best fits for a single species dimeric curve fit for N19a(cm) and a single species trimeric curve fit for L19a(cm). The upper panels show the residuals from the curve fits, and the lower panels show lnY versus r^2 of the data (filled circles) compared to theoretical monomer (m), dimer (d) and trimer (t) single species plots (brocken lines), where y is the fringe displacement, which is proportional to the concentration.

substituted at position 19**a** showed that the peptide sequence is an excellent model for probing the effects of single amino acid substitutions on oligomerization state in either the reduced or oxidized state.

The Asn analogs N19**a**(cm) and N19**a**(ox) were found to be exclusively two-stranded, almost certainly due to interchain H-bonding of the Asn side-chain, observed initially in dimeric DNA transcription regulatory coiled-coils (Harbury *et al.*, 1993; Lavigne *et al.*, 1997; Lumb & Kim, 1995; O'Shea *et al.*, 1991). It was noted that the N19**a**(ox) analog was exclusively monomeric, and did not form an extended two-stranded dimer (Figure VI-2b), presumably due to a single dimeric molecule with the same packing as two corresponding monomeric molecules (Figure VI-2a) being entropically disfavored. The proposed interhelical H-bond could occur between Asn residues on adjacent α-helices as seen in GCN-4 (O'Shea *et al.*, 1991), between Asn and the main chain carbonyl on the opposite chain as seen in the c-Myc-Max heterodimeric leucine zipper (Lavigne *et al.*, 1998) as well as between the substituted Asn side chains and Gln18**g** similar to that observed between Asp in a **g** position next to Asn in an **a** position of a α-tropomyosin/GCN4 hybrid coiled-coil (Greenfield *et al.*, 1998). It was interesting to note that while N19**a**(cm) was significantly less hydrophobic than the A19**a**(cm) analog as determined by C-8 reversed-phase HPLC ($\Delta t_{R(Ala)}$ -6.8 min), N19**a**(ox) was actually more stable than A19**a**(ox) ($\Delta\Delta G_{(Ala)}$ 0.77 kcal/mol). This untoward stabilization may well be attributable to the formation of such interhelical H-bonds.

While a Leu substitution could be expected to increase the stability of either a two- or three-stranded coiled-coil due to an increase in hydrophobicity, a three-stranded structure for L19**a** is clearly favored, due at least in part to position "**a**" side chains being ~7% more buried in a three-stranded coiled-coil than in a two-stranded coiled-coil (Gonzales *et al.*, 1996a). It is interesting to note that the I19**a** isostere was conformationally degenerate, folding into both two- and three-stranded structures, possibly due to the favorable packing of β-branched hydrophobic residues in the "**a**" position of two-stranded coiled-coils. Our results (Table VI-2) show that I19**a** was 0.3 kcal/mol more stable than L19**a**, with a GdnHCl midpoint 0.4 M higher. This may contribute to the two- and three-stranded conformers of I19**a** being closer in stability than the two- and three-stranded conformers of L19**a**, such that both are present in benign medium. It should be pointed out, however, that GCN4 analogs with either an $\mathrm{Ile_a}$-$\mathrm{Leu_d}$ or $\mathrm{Val_a}$-$\mathrm{Leu_d}$ hydrophobic core have been shown to form exclusively disulfide-bridged two-stranded coiled-coil monomers (Harbury *et al.*, 1993). These observations show that for some coiled-coils oligomerization state is not determined solely by positions "**a**" and "**d**". For example, positions "**e**" and "**g**", which are also involved in the hydrophobic interface, have been shown to play an important role (Alberti *et. al.*, 1993; Beck *et. al.*, 1997; Kammerer *et. al.*, 1998; Kohn *et. al.*, 1998; Zeng *et. al.*, 1997).

Quaternary structure determination

Sedimentation velocity experiments were performed on trimers L19**a**(ox) and L19**a**(red) to measure the axial ratio, (a/b), of these molecules. These experimentally obtained values were then used to evaluate the validity of their proposed quatenary structures, shown in Figures VI-2c and VI-2f, respectively, by comparison to the expected axial ratio values calculated using dimensions obtained from x-ray crystal structures of three-stranded coiled-coils of similar hydrophobic core structure to model X19**a** (Gonzales *et al.*, 1996a; Harbury *et al.*, 1994; Lovejoy *et al.*, 1993; Ogihara *et al.*, 1997) (Table VI-3). Thus, based on these published structures, the diameter of a three-stranded coiled-coil of peptide X19**a** is expected to be approximately 20 Å (b=10 Å), and the rise/residue is expected to be 1.5 Å. Reduced trimer L19**a**(red) was examined initially, and since this peptide sequence consists of 38 residues, the expected length is ~60 Å (a=30 Å). This results in an expected axial ratio a/b= ~3 for a coiled-coil of L19**a**(red). An axial ratio of 3.0 was in fact obtained experimentally for this analog, with a diameter of 24 Å and a length of 73Å (Table VI-3).

Table VI-3. Determination of peptide assymmetry by analytical ultacentrifugation.

Peptide	C_p (mg/ml) [a]	M_{cal} [b]	$\bar{v}$ (cm^3/g) [c]	$S^\circ_{20,w}$ (sec) [d]	f/f_0 [e]	a/b_{obs} [f]	a/b_{cal} [g]
L19a (ox)	2.5	22,812	0.7645	1.60	1.557	6.6	6
L19a (red)	1.5	11,409	0.7645	1.22	1.2846	3.0	3

[a] C_p is the peptide concentration.

[b] Calculated molecular weight of the trimer based on the amino acid sequence (Figure 1).

[c] $\bar{v}$ is the partial specific volume calculated using the method of Cohn and Edsall and the program Sednterp (Cohn & Edsall, 1943; Hayes *et al.*, 1998).

[d] Sedementation coefficient corrected to 20 °C in pure water and also corrected for solute concentration, calculated using the program Sednterp (Hayes *et al.*, 1998)

[e] Measure of the degree of assymmetry calculated from $S^\circ_{20,w}$ using the program Sednterp (Hayes *et al.*, 1998), where **f** is the translational frictional coefficient and f_0 is the calculated frictional coefficient of a sphere having the same volume as the protein molecule.

[f] Axial ratio modeled as a prolate ellipsoid (a = radius of the long dimension and b = radius of the short dimension of the ellipsoid) calculated from f/f_0 using the program Sednterp (Hayes *et al.*, 1998). For L19a(ox) length (2a) = 154 Å, diameter (2b) = 24 Å, and for L19a(red) 2a = 73 Å, 2b = 24 Å based on calculations from experimental data.

[g] Axial ratio calculated assuming L19**a**(ox) adopts the oxidized trimer quatenary structure and L19**a**(red) the reduced trimer quatenary structure depicted in figure 2, with diameter (2b) = 20Å. In the case of L19**a**(red) the length (2a) was calculated assuming 35 residues in an α-helix with a rise/residue of 1.5 Å for a total length of ~60 Å, and in the case of L19**a**(ox) the length (2a) was calculated assuming 70 residues in an α-helix with a rise/residue of 1.5 Å for a total length of ~120 Å.

We were interested in knowing if trimeric L19**a**(ox) forms an elongated three-stranded coiled-coil as in Figure VI-2c, or if the quatenary structure of the trimer consists of two-stranded oxidized molecules associated in a barrel-type arrangement. If the oxidized trimer L19**a**(ox) forms the three-stranded quaternary structure shown in Figure VI-2c, where each strand consists of two polypeptide chains linked by an interchain disulfide bridge, the length would be ~120 Å (a=60 Å), and the expected axial ratio a/b=~6 (Table VI-3). However, if the quaternary structure was actually a six-stranded barrel-type structure, composed of three oxidized two-stranded monomers packed side-to-side, the axis in the long dimension (a) would be the same length as seen for L19**a**(red). The axis in the shorter dimension (b) would increase, however, and the axial ratio (a/b) would therefore decrease from the value of 3.0 obtained for L19**a**(red). The experimentally determined axial ratio was in fact 6.6 for L19**a**(ox) (Table VI-3), with an a diameter of 24Å and a length of 154Å. This result clearly indicates a more elongated, asymmetric quaternary structure for L19**a**(ox) compared to L19**a**(red), and is entirely consistent with the proposed structure shown in Figure VI-2c. The excellent agreement between the experimentally obtained axial ratios and those calculated based on X-ray crystal structures clearly validates this method as a tool for assessing coiled-coil asymmetry and quatenary structure.

Relationship between protein concentration and stability

In order to obtain comparable stability data the unfolding process must occur from the same oligomerization state for all analogs. Therefore, the protein concentration dependence of stability as measured by $[\mathrm{GdnHCl}]_{1/2}[\theta]_{222}$, the concentration of GdnHCl required to give a 50% decrease in ellipticity at 222 nm, was determined. A change in transition midpoint with concentration is indicative of different oligomerization states prior to unfolding. It was observed that a change in oligomerization state actually occurs over the protein concentration range examined by performing HPSEC under benign

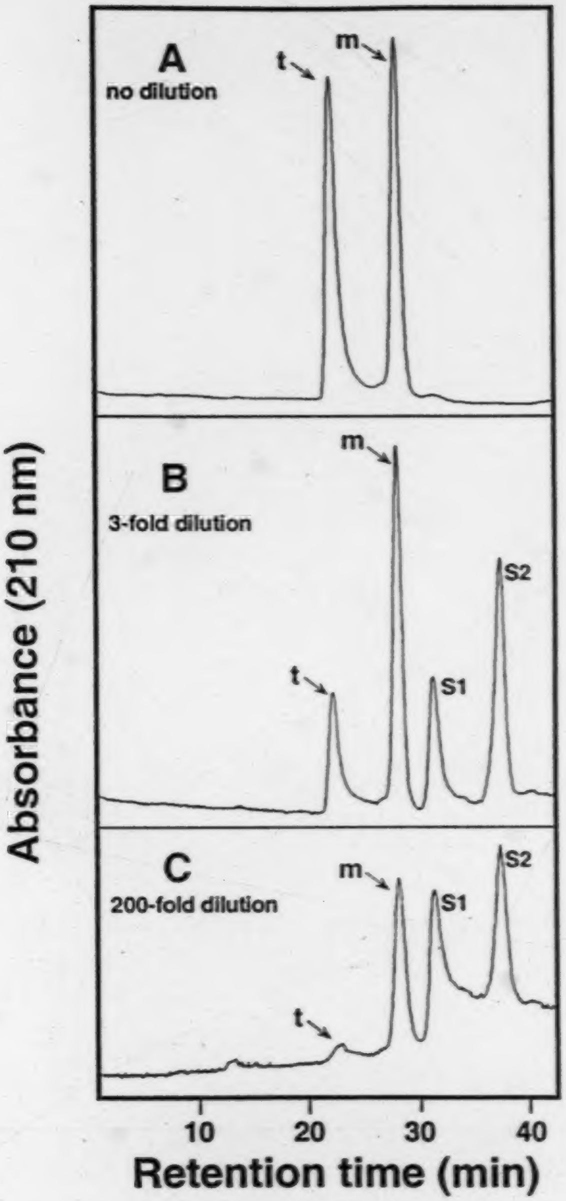


Figure VI-8. HPSEC of the I19a analog after equilibration at 500 µM (A), 170 µM (B) and 2.5 µM (C) prior to injection. S1 and S2 are peptide standards with sequences as shown in the materials and methods section. t, trimer; m, monomer.

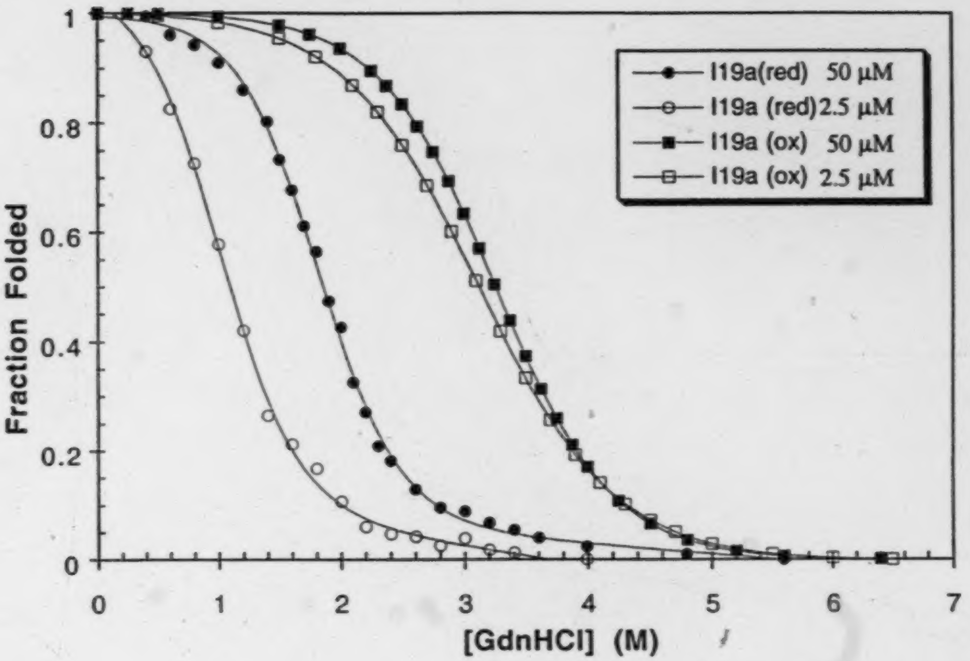


Figure VI-9. GdnHCl denaturation profiles of reduced (circles) and oxidized (squares) peptide I19a monitored by circular dichroism spectroscopy at 222 nm, and at peptide concentrations of 50 μM (filled symbols) and 2.5 μM (open symbols). Denaturation profiles were recorded at 25 °C in 100 mM KCl, 50 mM PO_4 pH 7.0 buffer. The fraction folded was calculated as described in the Materials and Methods section.

conditions on analogs I19**a**(ox), L19**a**(ox), and A19**a**(ox) at peptide concentrations ranging from 2.5 μM to 500 μM. At 500 μM I19**a**(ox) showed a nearly equimolar mixture of trimeric and monomeric species, which upon dilution to 2.5 μM was converted predominately to monomer (Figure VI-8, panels A-C). HPSEC analysis of 500 μM A19**a**(ox) revealed a trimer:monomer ratio of 8:2, which decreased to 3:2 upon dilution to 2.5 μM (HPSEC trace not shown). HPSEC analysis of 500 μM L19**a**(ox) revealed that this analog was >95% trimeric, and upon dilution to ~2.5 μM the amount of monomer increased to only ~20% of the total (HPSEC trace not shown). In the absence of an interhelical disulfide bridge, the denaturation midpoint of I19**a**(red) was expected to be peptide concentration dependent since the oligomerization state changed dramatically over the protein concentration range employed, and a substantial 0.8 M decrease in the denaturation midpoint was indeed observed (Figure VI-9). In contrast, for I19**a**(ox) and A19**a**(ox), in which the nascent α-helices are joined by a disulfide bridge, the GdnHCl midpoints showed little change when the peptide concentration was decreased from 500 μM to 2.5 μM, as shown in Figure VI-9 for I19**a**(ox), even though the oligomerization state changes from mainly trimeric at 500 μM to essentially monomeric at 2.5 μM in benign medium (Figure VI-8, panel C). This independence of the denaturation midpoint on peptide concentration is consistent with the oxidized trimeric coiled-coil species (Figure VI-2c) undergoing a silent transition (no detectable change in $[\theta]_{222}$) to a fully folded disulfide-bridged two-stranded coiled-coil monomer (Figure VI-2a) prior to unfolding and a loss of α-helical content. GdnHCl denaturation experiments were performed routinely at the lowest possible peptide concentration (2.5 μM) to further ensure that denaturation occurred from a disulfide-bridged two-stranded monomeric state.

HPSEC under denaturing conditions

In order to verify that denaturation was occurring from a monomeric two-stranded oligomerization state for all of the analogs, HPSEC was used to monitor the oligomerization state in the presence of increasing concentrations of GdnHCl prior to and through the denaturation transition, at peptide concentrations of 2.5 μM and GdnHCl concentrations ranging from 0.4 M to 4.0 M (Figures VI-10 and VI-11). Previous work in our laboratory employing HPSEC under different buffer conditions has shown that the size exclusion volume changes as the concentration of GdnHCl is increased (Mant & Hodges, 1991). Thus, in order for HPSEC analysis employing varying denaturant concentrations to be interpretable, it is imperative that molecular weight internal standards be present (Mant *et al*, 1987). The sequences of the peptides used as molecular weight standards in this study, denoted S1 and S2, were Ac-$(GKGLG)_6$-amide (30 residues, MW 2534 Da) and Ac-KYGLGGAGGLK-amide (11 residues, MW 1062 Da), respectively (Burke *et al.*, 1989; Mant *et al.*, 1997). Further, these standards were designed to be random coils under benign conditions (Mant & Hodges, 1991; Mant *et al.*, 1987), such that the relative retention times of S1 and S2 were then independent of GdnHCl concentration (Figures VI-10 and VI-11). The joint application of HPSEC and CD to the study of protein stability has been described previously (Becktel & Lindorfer, 1991; Corbett & Roche, 1984). This approach has the advantage of measuring completely different physical properties, α-helical content as measured by CD and hydrodynamic volume as measured by HPSEC, which can be useful in detecting stable intermediates. It was noted that the folded oxidized monomer and oxidized trimer eluted as distinct peaks, which indicates a slow exchange rate for interconversion of these oligomerization states relative to the elution time of ~60 min. The unfolded monomer peak (m_u) eluted well ahead of the folded monomer (m_f) (Figures VI-10 and VI-11), due to an increase in the hydrodynamic volume of the folded coiled-coil protein upon denaturation, analogous to that seen for denaturation of globular proteins (Stellwagen & Shalongo, 1991).

Figure VI-10, panels A-E, shows the change in HPSEC retention time of the N19**a**(ox) peptide that occurs during the GdnHCl denaturation process. The $[GdnHCl]_{1/2}[\theta]_{222}$ of N19**a**(ox) is 1.5 M (Table VI-2), and at GdnHCl concentrations of

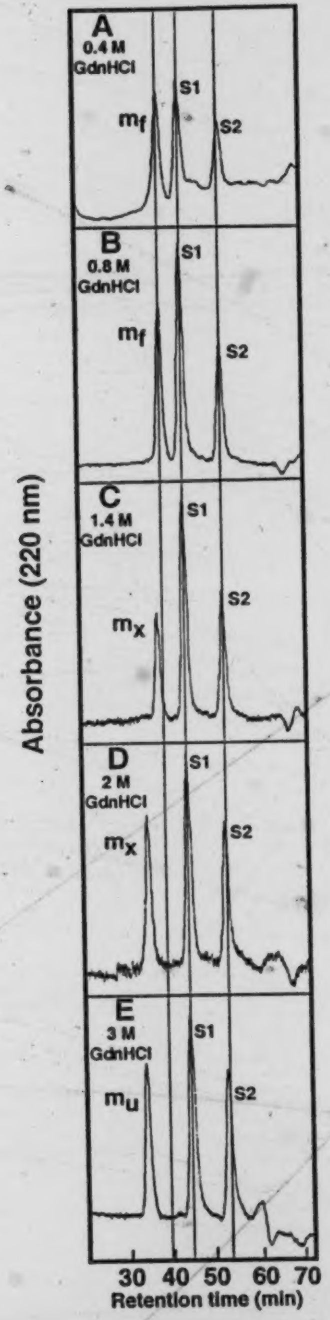


Figure VI-10. HPSEC of the N19a(ox) peptide at 2.5 µM under denaturing conditions ranging from 0.4 M GdnHCl (panel A) to 3 M GdnHCl (panel E). m_f, folded monomer; m_u, unfolded monomer; m_x, the peak resulting from rapid interconversion of m_f and m_u. S1 and S2 are random coil peptide standards described in the Materials and Methods section used to align the HPSEC traces due to changes in the column fractionation range with increasing GdnHCl concentration.

0.4 M and 0.8 M (Figure VI-10, panels A and B) N19**a**(ox) elutes as folded monomer m_f. This is consistent with the observation that at 0.8 M GdnHCl no significant change in ellipticity has yet occurred (Figure VI-5, top panel). As the GdnHCl concentration is increased to 1.4 M, 2.0 M and 3.0 M the retention time of the N19**a**(ox) peak, now designated as m_x, gradually decreases until the retention time of the unfolded monomer state m_u is reached at 3.0 M (Figure VI-10, panels C-E), after which at higher GdnHCl concentrations of 4.0 M and 5.0 M the relative retention time of m_u does not change (HPSEC traces not shown). The gradual shift in retention time of m_f towards that of m_u is consistent with fast exchange between m_f and m_u, such that the peak denoted m_x with an averaged retention time dependent on the ratio of the m_f and m_u states is observed. This is the first report of using HPSEC to monitor the denaturation of a folded protein in fast exchange with the corresponding unfolded state. It should be noted that the key to the successful use of this method is employment of random-coil peptide molecular weight internal standards.

Figure VI-11, panels A-E, shows the change in HPSEC retention time of the I19**a**(ox) peptide that occurs during the GdnHCl denaturation process. Analogs I19**a**(ox) and A19**a**(ox), which both populate a mixture of two- and three-stranded conformations under benign conditions, were each observed to undergo a shift towards the oxidized trimeric oligomerization state at the lowest GdnHCl concentration employed (0.4 M), shown in Figure VI-11 panel A for I19**a**(ox). Thomas and coworkers (1997) have recently reported a similar phenomenon occurring with their model peptide termed FZ, which was based on Hodges and coworkers' model peptide TM83 (Lau *et al.*, 1984). Similar to L19**a** in the present study, FZ forms trimeric coiled-coils even at low peptide concentrations, and moreover low concentrations of denaturant also promote further association of FZ. This formation of higher order oligomers in the presence of low denaturant concentrations may be due to partial unfolding or dissociation of the individual α-helices, which can allow rearrangement of structure. A further increase in GdnHCl concentration from 0.4 M

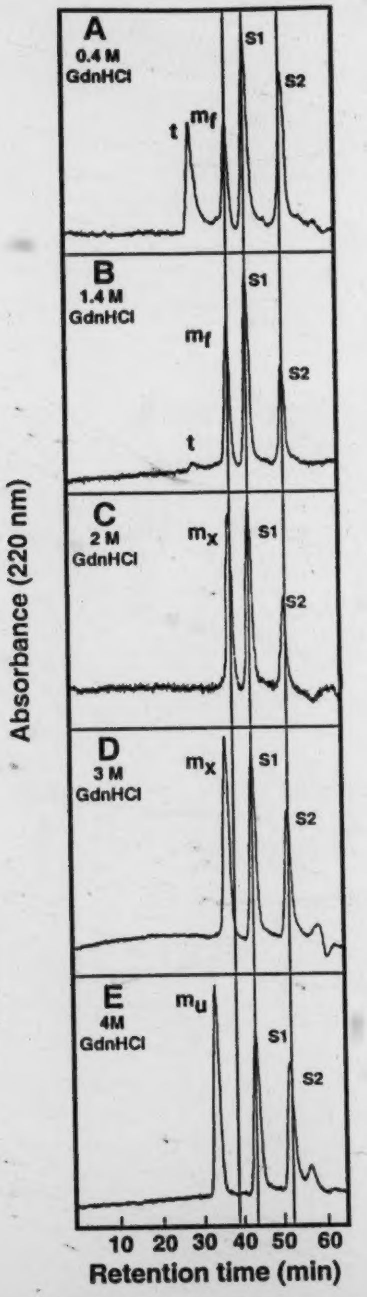


Figure VI-11. HPSEC of the I19a(ox) peptide at 2.5 μM under denaturing conditions ranging from 0.4 M GdnHCl (panel A) to 4 M GdnHCl (panel E). m_f, folded monomer; m_u, unfolded monomer; m_x, the peak resulting from rapid interconversion of m_f and m_u. S1 and S2 are random coil peptide standards described in the Materials and Methods section used to align the HPSEC traces due to changes in the column fractionation range with increasing GdnHCl concentration.

to 1.4 M caused a gradual shift in the trimer/monomer equilibrium toward there being virtually only folded monomeric species (m_f) by 1.4 M GdnHCl (Figure VI-11, panel B) before any change in α-helical content is observed by CD spectroscopy (Figure VI-5, panel A). This result corroborates the protein concentration independence of $[\text{GdnHCl}]_{1/2}[\theta]_{222}$ for analog I19**a**(ox), and proves that denaturation occurs from a two-stranded monomeric oligomerization state. In a previous study it was shown that a model coiled-coil that existed as an oxidized dimer under benign conditions was monomeric at a concentration of GdnHCl (2 M) well below the observed GdnHCl denaturation midpoint (5.1 M) in a transition that was silent to detection by CD spectroscopy at 222 nm, prior to unfolding from the disulfide-bridged monomeric state (Zhu *et al.*, 1993). The change in free energy associated with this equilibrium shift is equivalent to the free energy difference between packing the α-helices in a three-stranded coiled-coil and a two-stranded coiled-coil, which cannot be determined in this study because the transition is silent in CD spectroscopy. The denaturation of the folded monomeric state of I19**a**(ox) at 2 M GdnHCl and 3 M GdnHCl resulted in the gradual shift in retention time of peak m_x toward that of m_u (Figure VI-11, panels C and D), as seen previously for N19**a**(ox). At 4 M GdnHCl (Figure VI-11, panel E) the monomer is fully unfolded as expected based on the GdnHCl denaturation profile (Figure VI-5, top panel). For the L19**a**(ox) peptide, which is already mainly trimeric even at 2.5 μM in benign medium, a conversion to folded monomer m_f at higher GdnHCl concentrations was again observed, followed by denaturation to unfolded monomer m_u (HPSEC trace not shown).

Comparison of protein denaturation obtained from CD spectroscopy and HPSEC

The change in retention time of m_f, m_x, and m_u relative to the random coil peptide standard S1 was plotted against GdnHCl concentration (Figure VI-5, bottom panel), to allow calculation of $[\text{GdnHCl}]_{1/2}$HPSEC, the transition midpoint obtained from chemical denaturation experiments monitored by HPSEC. This value is defined as the concentration of GdnHCl required to shift the HPSEC retention time of m_x, which resulted from rapid equilibration of m_f and m_u, midway between the HPSEC retention times of m_f and m_u. Interestingly, the $[\text{GdnHCl}]_{1/2}$HPSEC values obtained for all the analogs tested were very similar to the $[\text{GdnHCl}]_{1/2}[\theta]_{222}$ values obtained by measuring the change in α-helical content by CD (Table VI-2 and Figure VI-5). The observation that the change in retention time of m_f via m_x towards the retention time of m_u is precisely mirrored by a concomitant loss in helicity as measured by CD indicates that the HPSEC retention time of m_x relative to m_f and m_u is a direct measure of the peptide fraction folded at a given GdnHCl concentration. Coincidence of the midpoints for chemical denaturation of myoglobin and T4 lysozyme measured by both CD and size-exclusion chromatography have been reported previously, and this is thought to be consistent with a simple two-state transition (Becktel & Lindorfer, 1991; Corbett & Roche, 1984). The combination of these techniques is particularly valuable in the present study since it allowed the direct observation that the unfolding process occurs from a folded disulfide-bridged two-stranded monomeric state regardless of the oligomerization state under benign conditions, thereby allowing measurement of relative stabilities resulting from single amino acid substitutions in a two-stranded, disulfide-bridged monomeric coiled-coil.

Oligomerization state control by disulfide bridge placement in the hydrophobic core

To further validate the importance of the flexibility of the Cys-Gly-Gly linker and to test whether oligomerization could be prevented in our model sequence, we changed the disulfide bond to position 36**d** in the hydrophobic core (C36**d** L19**a**, Figure VI-1). Hodges and coworkers had previously investigated the effect on stability and oligomerization state resulting from inserting an interhelical disulfide bond systematically in each "**a**" and "**d**" position of a five-heptad coiled-coil that contained Leu in each hydrophobic core position (Zhou *et al.*, 1993). They found that placing the disulfide bond at the N-terminal "**a**" position promoted the formation of higher order oligomers, while

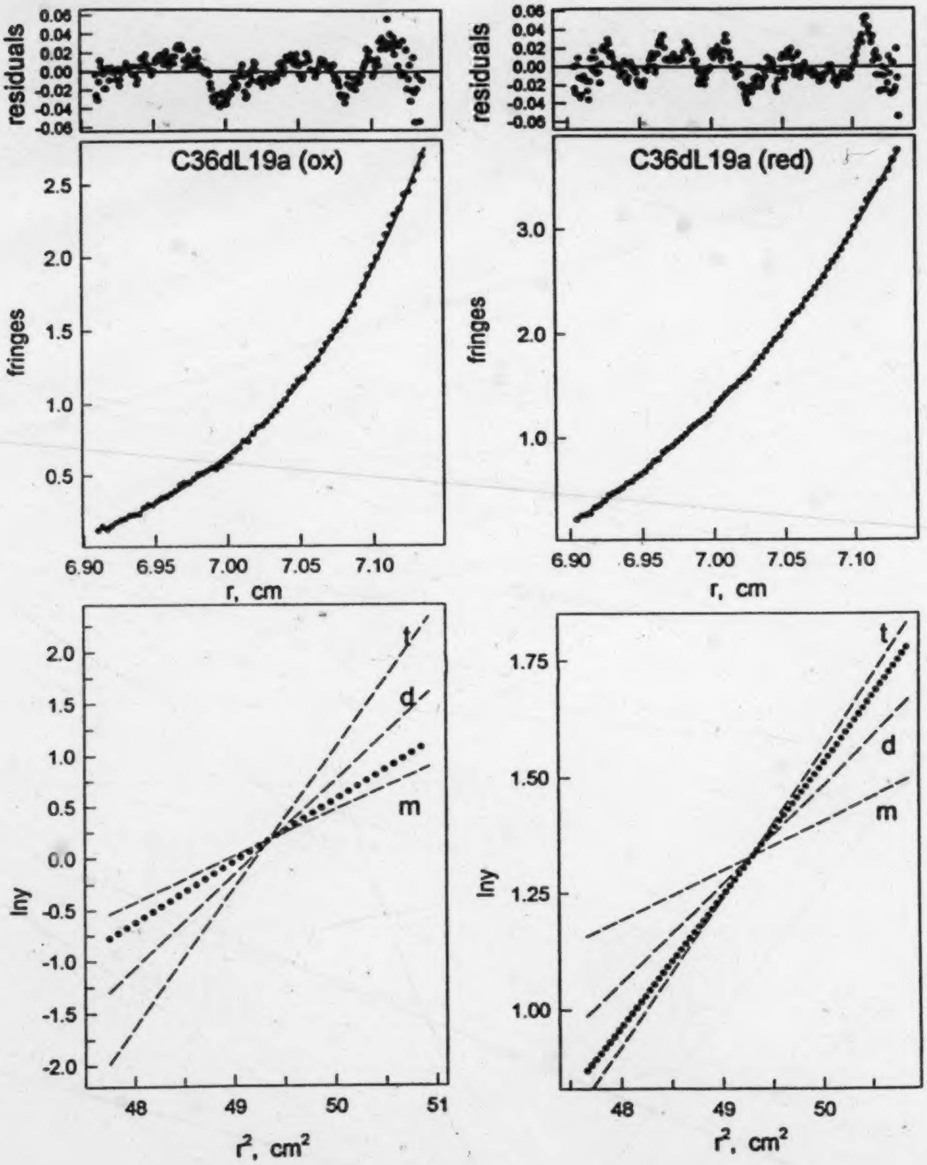


Figure VI-12. Sedimentation equilibrium analysis of C36dL19a(ox) and C36dL19a(red). The middle panels (fringes versus radius) show the best fits for a single species monomeric curve fit for C36dL19a(ox) and a single species trimeric curve fit for C36dL19a(red). The upper panels show the residuals from the curve fits, and the lower panels show lnY versus. r^2 of the data (filled circles) compared to theoretical monomer (m), dimer (d) and trimer (t) single species plots (dashed lines), where y is the fringe displacement, which is proportional to the concentration.

placing the disulfide bond at the C-terminal "**d**" position or any internal "**a**" or "**d**" position resulted in specifying a strictly monomeric conformation. In the present study, C36**d**L19**a**(red) approached a trimeric oligomerization state as determined by sedimentation equilibrium (Figure VI-12), likely because the sequence is an analog of the trimer L19**a**(red). However, C36**d**L19**a**(ox), which contains an interhelical disulfide bond at the C-terminal position, indeed forced a two-stranded conformation on the molecule (Figure VI-12). Results from HPSEC experiments in which an aliquot of C36**d**L19**a** eluted as a single peak with retention time coincident with that of folded oxidized monomer (HPSEC data not shown) supported this conclusion. It is possible that the three-stranded, disulfide-bridged trimeric quatenary structure with favorable side chain packing depicted in Figure VI-2c for the L19**a**(ox) parent sequence is energetically unfavorable for sequence C36**d**L19**a**(ox) due to steric hindrance for rotation about the C36 disulfide bond required to give the extended three-stranded coiled-coil. These results clearly show that using the C36 disulfide bond maintains this sequence in a two-stranded monomeric state under benign conditions. However, the Cys-Gly-Gly linker provides the advantage of flexibility such that single amino acid substitution can direct the oligomerization state.

C. Conclusions

In summary, we have designed the model peptide sequence X19**a** for use in determining the effect of single amino acid substitution in the hydrophobic core "**a**" position on coiled-coil hydrophobicity, stability and oligomerization state. The Cys-Gly-Gly flexible linker in model X19**a** allowed the intrinsic oligomerization state resulting from single amino acid substitution to be measured, while simultaneously limiting the GdnHCl denaturation profiles to be concentration independent due to denaturation from a disulfide-bridged two-stranded monomeric state. Interestingly, of these five analogs one specified a three-stranded coiled coiled-coil, one specified a two-stranded coiled-coil and the three other analogs showed mixtures of each. Thus, the sequence is sensitive to the effect various residues may have on oligomerization state, while allowing for the assessment of stability of each substituted residue in a two-stranded disulfide-bridged monomeric oligomerization state. In addition, we have established all the methods required for the biophysical characterization of each analog, which includes a novel HPSEC procedure for monitoring unfolding where the unfolded and folded states are in fast exchange. Further studies are in progress that will provide a complete data set for the quantitative effect on hydrophobicity, stability and oligomerization state resulting from substituting all naturally occurring amino acids in model peptide sequence X19**a**.

D. References

Adamson, J. G., Zhou, N. E. & Hodges, R. S. (1993). Structure, function and application of the coiled-coil protein folding motif. *Curr. Opin. Biotech.* **4,** 428-437.

Alber, T. (1992). Structure of the leucine zipper. *Curr. Opin. Genet. Dev.* **2,** 205-210.

Alberti, S., Oehler, S., von Wilcken-Bergmann, B.& Müller-Hill, B. (1993). Genetic analysis of the leucine heptad repeats of Lac repressor: Evidence for a 4-helical bundle. *EMBO J.* **12,** 3227-3236.

Beck, K., Gambee, J. E., Kamawal, A.& Bächinger, H. P. (1997). A single amino acid can switch the oligomerization state of the α-helical coiled-coil domain of the cartilage matrix protein. *EMBO J.* **16,** 3767-3777.

Beckman Instruments, Inc. (1997a). Optima XL-I Analytical Ultracentrifuge Manual. Spinco Business Center of Beckman Instruments, Inc. Palo Alto, CA.

Beckman Instruments, Inc. (1997b). Optima XL-A/ XL-I Data Analysis Software, Version 4.0., copyright (c) 1997. Spinco Business Center of Beckman Instruments, Inc. Palo Alto, CA.

Becktel, W. J. & Lindorfer, M. A. (1991). Complementary application of size-exclusion chromatography and circular dichroism to the study of protein folding. In *High-Performance Liquid Chromatography of Peptides and Proteins: Separation, Analysis, and Conformation* (Mant, C. T. & Hodges, R.S., eds.), pp. 605-611, CRC Press, Boca Raton, FL.

Betz, S., Fairman, R., O'Neil, K., Lear, J. & DeGrado, W. (1995). Design of two-stranded and three-stranded coiled-coil peptides. *Phil. Trans. R. Soc. London B* **348,** 81-88.

Burke, T. W. L., Mant, C. T., Black, J. A. & Hodges, R. S. (1989). Strong cation-exchange high-performance liquid chromatography of peptides: Effect of non-specific hydrophobic interactions and linearization of peptide retention behaviour. *J. Chromatogr.* **476,** 377-389.

Cedervall, T., Johansson, M. U. & Åkerström, B. (1997). Coiled-coil structure of group A streptococcal proteins. Different temperature stability of class A and C proteins by hydrophobic-nonhydrophobic amino acid substitution a heptad positions a and d. *Biochemistry* **36,** 4987-4994.

Chen, Y.-H., Yang, J. T. & Chau, K. H. (1974). Determination of the helix and β form of proteins in aqueous solution by circular dichroism. *Biochemistry* **13,** 3350-3359.

Chervenka, C. H. (1970). A Manual of Methods for the Analytical Ultracentrifuge. Spinco Business Center of Beckman Instruments, Inc. Palo Alto, CA.

Chong, P. C. S. & Hodges, R. S. (1982). Proximity of sulfhydryl groups to the sites of interaction between components of the troponin complex from rabbit skeletal muscle. *J. Biol. Chem.* **257,** 2549-2555.

Cohen, C.& Parry, D. A. D. (1990). α-helical coiled coils and bundles: How to design an α-helical protein. *Proteins: Struct. Funct. Genet.* **7,** 1-15.

Cohn, E. J. & Edsall, J. T. (1943). Proteins, Amino Acids and Peptides as Ions and Dipolar Ions, pp. 370-381, Reinhold, New York.

Cooper, T. M. & Woody, R. W. (1990). The effect of conformation on the CD of interacting helices: A theoretical study of tropomyosin. *Biopolymers* **30,** 657-676.

Corbett, R. J. T. & Roche, R. S. (1984). Use of high-speed size-exclusion chromatography for the study of protein folding and stability. *Biochemistry* **23,** 1888-1894.

Crick, F. H. C. (1953). The packing of α-helices: Simple coiled-coils. *Acta Crystallogr.* **6,** 689-698.

Dill, K. A., Bromberg, S., Yue, K., Fiebig, K. M., Yee, D. P., Thomas, P. D. & Chan, H. S. (1995). Principles of protein folding-A perspective from simple exact models. *Protein Sci.* **4,** 561-602.

Gonzales, L. Jr., Woolfson, D. N. & Alber, T. (1996a). Buried polar residues and structural specificity in the GCN4 leucine zipper. *Nature Struct. Biol.* **3,** 1011-1018.

Gonzales, L. Jr., Brown, R. A., Richardson, D. & Alber, T. (1996b). Crystal structures of a single coiled-coil peptide in two oligomeric states reveal the basis for structural polymorphism. *Nature Struct. Biol.* **3,** 1002-1009.

Greenfield, N. J. & Hitchcock-DeGregori, S. E. (1995). The stability of tropomyosin, a two-stranded coiled-coil protein, is primarily a function of the hydrophobicity of residues at the helix-helix interface. *Biochemistry* **34,** 16797-16805.

Greenfield, N. J., Montelione, G. T., Farid, R. S. & Hitchcock-DeGregori, S. E. (1998). The structure of the N-terminus of striated muscle α-tropomyosin in a chimeric peptide: nuclear magnetic resonance structure and circular dichroism studies. *Biochemistry* **37,** 7834-7843.

Harbury, P. B., Zhang, T., Kim, P. S. & Alber, T. (1993). A Switch between two-, three-, and four-stranded coiled-coils in GCN4 leucine zipper mutants. *Science* **262,** 1401-1407.

Harbury, P. B., Kim, P. S. & Alber, T. (1994). Crystal structure of an isoleucine-zipper trimer. *Nature* **371,** 80-83.

Hayes, D. B., Laue, T. & Philo, J. (1998). Sedimentation Interpretation Program, Version 1.01. Copyright (c) 1995-1998.

Hodges, R. S. (1992). Unzipping the secrets of coiled-coils. *Current Biology* **2,** 122-124.

Hodges, R. S. (1996). De novo design of α-helical proteins: Basic research to medical applications. *Biochem. Cell Biol.* **74,** 133-154.

Hodges, R. S., Sodek, J., Smillie, L. B. & Jurasek, L. (1972). Tropomyosin: Amino acid sequence and coiled-coil structure. *Cold Spring Harbor Symp. Quant. Biol.* **37,** 299-310.

Hu, J. C. & Sauer, R. T. (1992). The basic-region leucine-zipper family of DNA binding proteins. *Nucleic Acids Mol. Biol.* **6,** 82-101.

Johnson, M. L., Correia, J. J., Yphantis, D. A. & Halvorson, H. R. (1981). Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques. *Biophys. J.* **36,** 575-588.

Kammerer, R. A. (1997). α-Helical coiled-coil oligomerization domains in extracellular proteins. *Matrix Biol.* **15,** 555-565.

Kammerer, R. A., Schulthess, T., Landwehr, R., Lustig, A., Fischer, D.& Engel, J. (1998). Tenascin-C hexabrachion assembly is a sequential two-step process initiated by coiled-coil α-helices. *J. Biol. Chem.* **273,** 10602-10608.

Kohn, W. D.& Hodges, R. S. (1998). *De novo* design of α-helical coiled-coils and bundles: Models for the development of protein-design principles. *Trends Biotech.* **16,** 379-389.

Kohn, W. D., Kay, C. M. and Hodges, R. S. (1995). Protein destabilization by electrostatic repulsions in the two-stranded α-helical coiled-coil. *Protein Sci.* **4,** 237-250.

Kohn, W. D., Kay, C. M. and Hodges, R. S. (1998). Orientation, positional, additivity and oligomerization state effects of interhelical ion pairs in α-helical coiled-coils. *J. Mol. Biol.* **283**, 993-1012.

Lau, S. Y. M., Taneja, A. K. & Hodges, R. S. (1984). Synthesis of a model protein of defined secondary and quaternary structure: Effect of chain length on the stabilization and formation of two-stranded α-helical coiled-coils. *J. Biol. Chem.* **259,** 13253-13261.

Lavigne, P., Crump, M. P., Cagné, S. M., Sykes, B. D., Hodges, R. S. & Kay, C. M. (1997). ^{1}H-NMR evidence for two buried Asn side-chains in the c-Myc-Max heterodimeric α-helical coiled-coil. Tech. Prot. Chem. **8,** 617-624.

Lavigne, P., Crumo, M. P., Cagné, S. M., Hodges, R. S., Kay, C. M. & Sykes, B. D. (1998). Insights into the mechanism of heterodimerization from the ^{1}H-NMR solution structure of the c-Myc-Max heterodimeric leucine zipper. *J. Mol. Biol.* **281,** 165-181.

Lovejoy, B., Choe, S., Cascio, D., McRorie, D. K., DeGrado, W. F. & Eisenberg, D. (1993). Crystal structure of a synthetic triple-stranded α-helical bundle. *Science* **259,** 1288-1293.

Lumb, K. J. & Kim, P. S. (1995). A buried polar interaction imparts structural uniqueness in a designed heterodimeric coiled-coil. *Biochemistry* **34,** 8642-8648.

Lupas, A. (1996). Coiled-coils: New structures and new functions. *Trends Biochem. Res.* **21,** 375-382.

Mant, C. T. & Hodges, R. S. (1987). Monitoring free silanols on reversed-phase supports with peptide standards. *Chromatographica* **24,** 805-814.

Mant, C. T. & Hodges, R. S. (1991). Requirement for peptide standards to monitor ideal and non-ideal behavior in size-exclusion chromatography. In *High-Performance Liquid Chromatography of Peptides and Proteins: Separation, Analysis, and Conformation* (Mant, C. T. & Hodges, R.S., eds), pp. 125-134, CRC Press, Boca Raton, FL.

Mant, C. T., Parker, J. M. R.& Hodges, R. S. (1987). Size-exclusion high-performance liquid chromatography of peptides: Requirement for peptide standards to monitor column performance and non-ideal behaviour. *J. Chromatogr.* **397,** 99-112.

Mant, C. T., Chao, H. & Hodges, R. S. (1997). Effect of mobile phase on the oligomerization state of α-helical coiled-coil peptides during high-performance size-exclusion chromatography. *J. Chromatogr. A* **791,** 85-98.

Monera, O. D., Sereda, T. J., Zhou, N. E., Kay, C. M. & Hodges, R. S. (1995). Relationship of sidechain hydrophobicity and α-helical propensity on the stability of the single-stranded amphipathic α-helix. *J. Peptide Sci.* **1,** 319-329.

Monera, O. D., Sönnichsen, F. D., Hicks, L., Kay, C. M. & Hodges, R. S. (1996). The relative positions of alanine resid[illegible] in the hydrophobic core control the formation of two-stranded or four-stranded α-helical coiled-coils. *Protein Eng.* **9,** 353-363.

O'Shea, E. K., Klemm, J. D., Kim, P. S. & Alber, T. (1991). X-ray structure of the GCN4 leucine Zipper, a two-stranded, parallel coiled-coil. *Science* **254,** 539-544.

Ogihara, N. L., Weiss, M. S., DeGrado, W. F. & Eisenberg, D. (1997). The crystal structure of the designed trimeric coiled-coil coil-V_aL_d: Implications for engineering crystals and supramolecular assemblies. *Protein Sci.* **6,** 80-88.

Pace, C. N. (1986). Determination and analysis of urea and guanidinium hydrochloride denaturation curves. *Methods Enzymol.* **131,** 266-280.

Sali, D., Baycroft, M. & Fersht, A. R. (1991). Surface electrostatic interactions contribute little to stability of barnase. *J. Mol. Biol.* **220,** 779-788.

Shoemaker, K. R., Kim, P. S., York, E. J., Stewart, J. M. & Baldwin, R. L. (1987). Tests of the helix dipole model for stabilization of α-helices. *Nature* **326,** 563-567.

Sodek, J., Hodges, R. S., Smillie, L. B. & Jurasek, L. (1972). Amino acid sequence of rabbit skeletal tropomyosin and its coiled-coil structure. *Proc. Natl. Acad. Sci* **69,** 3800-3804.

Sönnichsen, F. D., Van Eyk, J. E., Hodges, R. S. & Sykes, B. D. (1992). Effect of trifluoroethanol on protein secondary structure: An NMR and CD study using a synthetic actin peptide. *Biochemistry* **31,** 8790-8798.

Stellwagen, E. & Shalongo, W. (1991). Analysis of protein folding by size-exclusion chromatography. In *High-performance Liquid Chromatography of Peptides and Proteins: Separation, Analysis, and Conformation* (Mant, C. T. & Hodges, R.S., eds), pp. 599-604, CRC Press, Boca Raton, FL.

Thomas, R. M., Zampieri, A., Jumel, K. & Harding, S. E. (1997). A trimeric, alpha-helical, coiled-coil peptide: Association stoichiometry and interaction strength by analytical ultracentrifugation. *Eur. Biophys. J.* **25,** 405-410.

Wilson, I. A., Skehel, J. J. & Wiley, D. C. (1981). Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* **289,** 366-373.

Yphantis, D. A., (1991). Nonlinear Least Squares Program for Analysis of Equilibrium Ultracentrifugation Experiments. Copyright (c) 1991 by David A. Yphantis. 99 River Rd., Mansfield Center, Connecticut, U.S.A. 06259-1018.

Zeng, X., Zhu, H., Lashuel, H. A.& Hu, J. C. (1997). Oligomerization properties of GCN4 leucine zipper **e** and **g** position mutants. *Prot. Sci.* **6,** 2218-2226.

Zhou, N. E., Kay, C. M. & Hodges, R. S. (1992a). Synthetic model proteins: The relative contribution of leucine residues at the nonequivalent positions of the 3-4 hydrophobic repeat to the stability of the two-stranded α-helical coiled-coil. *Biochemistry* **31,** 5739-5746.

Zhou, N. E., Kay, C. M. & Hodges, R. S. (1992b). Synthetic model proteins: Positional effects of interchain hydrophobic interactions on stability of two-stranded α-helical coiled-coils. *J. Biol. Chem.* **267,** 2664-2670.

Zhou, N. E., Kay, C. M. & Hodges, R. S. (1993). Disulfide bond contribution to protein stability: Positional effects of substitution in the hydrophobic core of the two-stranded α-helical coiled-coil. *Biochemistry* **32,** 3178-3187.

Zhou, N. E., Kay, C. M. & Hodges, R. S. (1994). The net energetic contribution of interhelical electrostatic attractions to coiled-coil stability. *Protein Eng.* **7,** 1365-1372.

Zhu, B.-Y., Zhou, N. E., Kay, C. M. & Hodges, R. S. (1993). Packing and hydrophobicity effects on protein folding and stability: Effects of β-branched amino acids, valine and isoleucine, on the formation and stability of two-stranded α-helical coiled-coils/leucine zippers. *Protein Sci.* **2,** 383-394.

Chapter VII

The role of position "a" in determining the stability and oligomerization state of α-helical coiled-coils: 20 amino acid stability coefficients in the hydrophobic core of proteins

A version of this chapter has been submitted for publication: Kurt Wagschal, Brian Tripet, Pierre Lavigne, Colin Mant and Robert S. Hodges, *Protein Science* (1999).

A. Introduction

The α–helical coiled-coil, a left-handed superhelix of two or more right-handed α-helices first proposed by Crick (1), consists of a regularly repeating heptad unit designated *abcdefg* in which positions "**a**" and "**d**" are occupied predominately by hydrophobic amino acids, and positions "**e**" and "**g**" by charged residues. The repeating pattern of hydrophobic residues in the "**a**" and "**d**" positions creates an amphipathic face where the helices pack together at an angle of ~20° with interactions "**a**" to "**a'**" and "**d**" to "**d'**". Residues in the "**e**" and "**g**" positions on adjacent helices are often involved in interhelical salt bridges with interactions *i* to *i'* + 5 or "**g**" to "**e'**". In nature the coiled-coil motif has been elegantly employed to perform a diverse array of structural and dynamic functions, which include cytoskeletal proteins, DNA transcription factors and pH sensitive hinges that control the release of viral particles into cells (for recent reviews see 2-12). Recent inspection of the Brookhaven Protein Database has revealed that over 50 high-resolution structures of coiled-coil domains have been solved. Coiled-coils have been predicted to occur in well over 200 proteins, and will certainly be found in as yet undiscovered protein structures. It is clear, then, that identifying the occurrence of the coiled-coil motif and the structural features that modulate its overall stability and oligomerization state is vital to our understanding of a diverse array of phenomena in both normal and diseased cells, as well as being important for *de novo* protein design applications and our general understanding of protein folding and stability (3).

In either native sequence analysis or *de novo* protein design, the ability to predict the stability as well as the oligomerization state of coiled-coils from naturally occurring or designed protein sequences is paramount. However, no systematic study has yet been carried out to determine the contribution each side chain can make, when in the hydrophobic core, towards stability and oligomerization state. To this end, it has been established that while the factors important for stabilizing coiled-coils include the α-helical propensity of the residues and the presence or absence of intrachain and interchain salt bridges, it is the nature of the hydrophobic core "**a**" and "**d**" positions that provides the greatest contribution to overall stability, as well as being important in determining oligomerization state (13-18). As such, determining the relative contribution of each amino acid side chain toward overall stability when substituted in the hydrophobic core "**a**" and "**d**" positions would be extremely useful in predicting the occurrence of a coiled-coil from amino acid sequence, oligomerization state, orientation of polypeptide chains (parallel or antiparallel), preference for homodimerization or heterodimerization and overall stability. Algorithms designed to predict formation and oligomerization state of coiled-coil sequences based on the frequency of residue occurrence in a coiled-coil sequence database have been developed (19-23). However, frequency of occurrence of a given amino acid side chain in a hydrophobic core position of coiled-coils does not necessarily correlate with the thermodynamic stability imparted by the side chain, since a given residue can be preferred owing to the structural uniqueness engendered by the substitution, often at the expense of overall thermodynamic stability (13, 24, 25). This clearly renders a relative stability scale for amino acid side chain substitution in the hydrophobic core based on the frequency of residue occurrence untenable.

Examination of naturally occurring coiled-coils has revealed that all naturally occurring amino acid side chains excluding Pro have been used by nature in the hydrophobic core of coiled-coils, presumably to control overall stability as well as oligomerization state in the structural and functional context in which they are found. Hodges and coworkers designed the first model coiled-coil protein and demonstrated its utility in studying *de novo* design principles related to protein folding and stability (26), and this has been followed by extensive work on the use of coiled-coils to investigate principles of protein design and for biotechnology applications, recently reviewed (3, 8, 12, 27). Recently, we reported the *de novo* design of a model coiled-coil protein termed X19**a** (28), which consists of two identical 38-residue (5 heptad) amphipathic helices containing the 3-4 hydrophobic repeat characteristic of coiled-coils connected by a flexible Cys-Gly-Gly linker (Figure VII-1). Our goal was the development of a model host-guest peptide sequence able to measure the relative contribution toward overall coiled-coil stability and oligomerization state of 20 different amino acid substitutions at position ***a*** in the hydrophobic core. Five strategically selected analogs of this model peptide sequence were prepared to determine all of the biophysical techniques that would be required to fully characterize the coiled-coils. Preliminary results suggested that this peptide sequence (Figure VII-1) could indeed accept the full range of both stabilizing, hydrophobic substitutions as well as potentially destabilizing, hydrophilic substitutions that would be expected upon substitution with the naturally occurring amino acids (28).

In the present study, we have substituted 14 additional naturally occurring amino acids (cysteine was omitted and ornithine was included) to obtain a complete data set for the role of side-chain hydrophobicity in the hydrophobic core on protein stability of two-stranded coiled-coils. In addition, sedimentation equilibrium ultracentrifugation coupled with the results from HPSEC allowed us to assess the effect of amino acid substitution in position "**a**" of the central heptad of a coiled-coil on resultant oligomerization state. Finally, we present initial modeling results which rationalizes the observation of a unique oligomerization state for some analogs in terms of side-chain packing in the hydrophobic core. This is the first study to provide the effect of 20 different amino acid side-chains in the hydrophobic core of a coiled-coil protein on stability and oligomerization state. These results will be extremely important in the *de novo* design of proteins and in predicting coiled coil structure and stability in native proteins.

B. Results and Discussion

Peptide design

The *de novo* design and complete biophysical characterization of the model peptide sequence termed X19**a** (Figure VII-1) has been described previously in chapter VI where five strategically selected analogs were synthesized to test the utility of the sequence for the present study in terms of overall stability and oligomerization state, and further to ascertain the biophysical methods required to fully characterize the structure, stability and oligomerization state of the coiled-coils (28). Briefly, the heptad sequences used in the design were based on the amino acid sequence EVGALKK used in heptads 4 and 5. In heptads 1 and 2 a Lys residue was changed to an Ala residue at an "**f**" position to control overall charge of the peptide at neutral pH. Heptads 1, 4 and 5 provide a Val$_{\mathbf{a}}$-Leu$_{\mathbf{d}}$ hydrophobic core as well as optimal **e-g'** salt bridges that help ensure parallel, in-register α-helix alignment(29). The site of substitution is at position 19**a** in the central heptad, denoted with an "**X**" in Figure VII-1. Gln residues replaced the Glu and Lys residues at the "**e**" and "**g**" positions in both heptads 2 and 3 to prevent interaction of charged groups with the substituted amino acid side-chain. Also, Cys-Gly-Gly was added to NH_2-terminus (30) to permit disulfide bond formation and to provide a flexible linker that allowed unconstrained helical alignment. This flexible linker proved to be critical since it allowed an oxidized, three-stranded trimeric state (Figure VII-2b) to be observed in benign medium for some analogs. On the other hand, even though we observed a trimeric state for some of the analogs, during chemical denaturation we observed

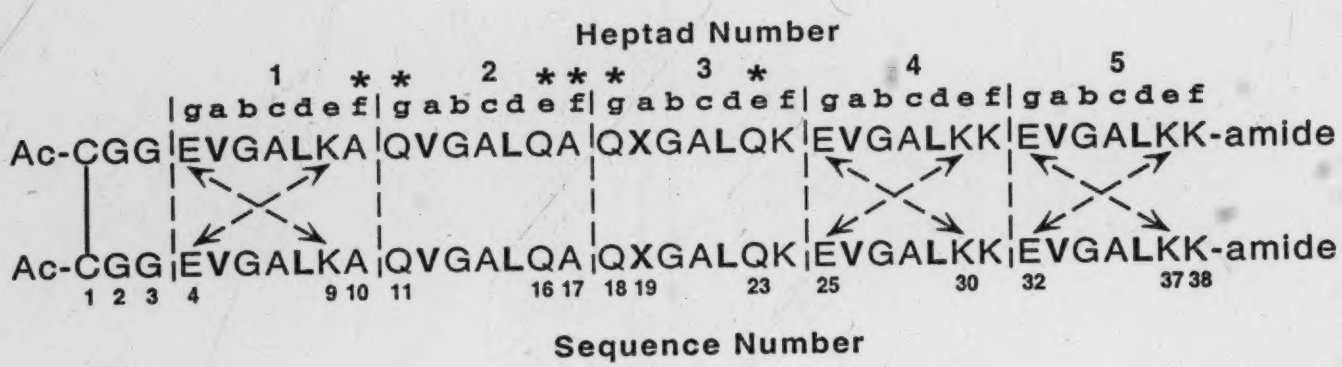

Heptad Number
1 * * 2 * * * 3 * 4 5
|g a b c d e f| g a b c d e f| g a b c d e f| g a b c d e f| g a b c d e f
Ac-CGG EVGALKA QVGALQA QXGALQK EVGALKK EVGALKK-amide
Ac-CGG EVGALKA QVGALQA QXGALQK EVGALKK EVGALKK-amide
1 2 3 4 9 10 11 16 17 18 19 23 25 30 32 37 38
Sequence Number

Figure VII-1. Amino acid sequence of the 76-residue disulfide bridged model coiled-coil X19a. Each polypeptide chain consists of 5 heptads and the linker region (Cys-Gly-Gly). The site of amino acid substitution is denoted "**X**" at position 19a; thus, the Leu analog is denoted L19a and the Val analog is denoted V19a, etc. Vertical dashed lines separate the heptads, and dashed arrows denote interchain salt bridges in heptads 1, 4, and 5. The model heptad sequence EVGALKK was employed for heptads 4 and 5, while changes were made in the model heptad sequence in heptads 1, 2 and 3, noted by "*". To avoid intrachain i to i+3 or i to i+4 charge-charge interactions with the site of substitution "X", charged residues involved in **e-g'** salt bridges in the standard heptad were replaced by Gln (positions 11, 16, 18 and 23). To adjust the overall charge at pH 7.0, Lys residues 10 and 17 in the standard heptad were replaced with Ala.

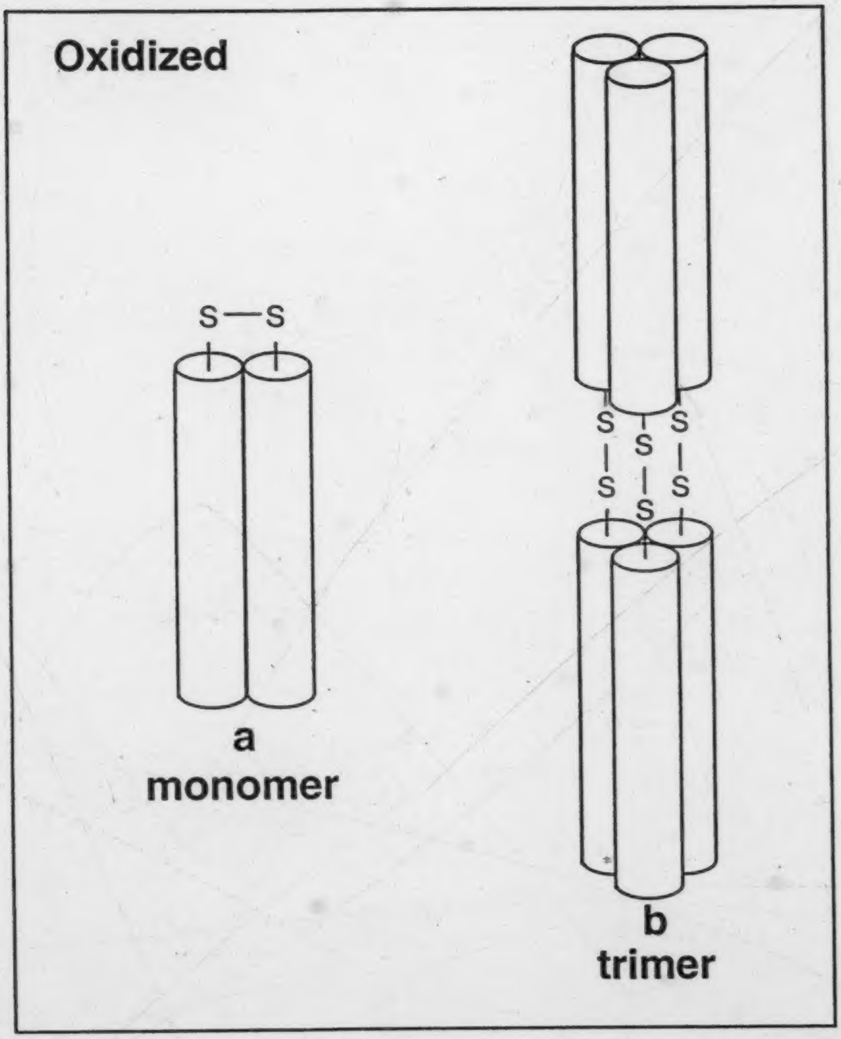


Figure VII-2. Proposed models of the interchain disulfide-bridged peptides Axial ratios obtained from sedimentation velocity experiments (28) indicate the monomer (a) is a two-stranded coiled-coil and the trimer (b) is an extended three-stranded coiled-coil. The cylinders represent the 35-residue polypeptide chain folded as an α-helix.

conversion of oxidized trimers to two-stranded oxidized monomers (Figure VII-2a) prior to denaturation to the unfolded state and any loss of helicity as judged by CD (28). As a result, the stability data for the entire analog set described herein can be compared since each X19**a** analog is unfolding from the same disulfide-bridged two-stranded monomeric oligomerization state, to a similar unfolded monomeric state.

CD spectroscopy

The CD spectra of each oxidized analog was measured under physiologically relevant, benign buffer conditions (100 mM KCl, 50 mM PO_4, pH 7.0), and also in 50% TFE at both 25 °C and 4.8 °C for all analogs (Table VII-1). The estimated maximal theoretical molar ellipticity at 222 nm and 25 °C is -34,730 deg•cm^2•dmol^{-1} for the 76-residue oxidized analogs (31, 32). The analogs generally approached this maximal value at 25 °C under benign conditions except for Ser19**a**, Arg19**a**, Gly19**a**, Orn19**a**, Asp 19**a** and Pro19**a** which were 94%, 78%, 79%, 84%, 18% and 5% helical, respectively, at 25 °C. These lower values indicate that the peptide populations of these analogs were either not completely in the folded state ($P_f \neq 1$), or in the case of the latter two peptides were essentially in the unfolded state, under benign conditions at room temperature. Some of the analogs were also examined at 4.8 °C (Table VII-1), and the molar ellipticity generally did not appreciably change for those analogs tested which were completely in the folded state at 25 °C (Asn, Ala, Gln, His, Glu and Lys). Among those analogs that were not completely in the folded state ($P_f \neq 1$) at 25 °C, the molar ellipticity increased by ~10% for the Ser and Arg analogs, while Gly19**a** and Orn19**a** showed a marked increase in molar ellipticity, indicating that these analogs were completely in the folded state at the lower temperature. It was especially interesting to note that Asp19**a**, which was essentially unfolded at 25 °C (-6,200 deg•cm^2•dmol^{-1}) became highly helical at 4.8 °C (-32,300 deg•cm^2•dmol^{-1}). Thus, at 4.8 °C all analogs except Pro19**a** were highly α-helical, which

Table VII-1. CD Spectroscopy Data

Substituted Amino Acid[a]	$[\Theta]_{222}$ [b] 25 °C Benign	% α-Helix [c] 25 °C Benign	$[\Theta]_{222}$ [b] 4.8 °C Benign	$[\Theta]_{222}$ [d] 25 °C 50%TFE	$[[\Theta]_{222/208}$ [e] 25 °C Benign	$[[\Theta]_{222/208}$ [e] 4.8 °C Benign	$[[\Theta]_{222/208}$ [e] 25 °C 50%TFE
Ile	-36100	104	-	-30800	1.06	-	0.90
Val	-38200	110	-	-33100	1.03	-	0.90
Met	-34800	100	-	-33600	1.06	-	0.89
Leu	-36600	105	-	-33300	1.03	-	0.89
Phe	-34300	99	-	-30200	1.04	-	0.88
Tyr	-36700	106	-	-33200	1.05	-	0.89
Trp	-39900	115	-	-32400	1.04	-	0.84
Asn	-34500	99	-33500	-33600	1.03	1.07	0.91
Thr	-39300	116	-	-28900	1.06	-	0.85
Ala	-34800	100	-36400	-30200	1.08	1.11	0.85
Gln	-35300	102	-33500	-29300	1.04	1.08	0.85
His	-37200	107	-36700	-32800	1.04	1.08	0.85
Lys	-37600	108	-36800	-28400	1.02	1.06	0.83
Ser	-32800	94	-36700	-29300	1.03	1.07	0.84
Arg	-27100	78	-30300	-22600	1.00	1.03	0.82
Gly	-27600	79	-35800	-28300	1.00	1.09	0.82
Orn	-29100	84	-36300	-31000	0.96	1.05	0.83
Glu	-35900	103	-38000	-25300	1.07	1.10	0.83
Asp	**-6200**	18	**-32300**	-24500	0.56	1.05	0.80
Pro	**-1900**	5	**-1900**	-20300	0.24	0.24	0.77

Table VII-1 Legend

---

[a] Substituted amino acid at position 19**a** denotes the corresponding disulfide bridged coiled-coil. See Figure VII-1 for peptide amino acid sequence.

[b] Mean residue molar ellipticities at 222 nm (deg•cm^2•dmol^{-1}). Measurements were performed on ~50 µM samples of oxidized (disulfide-bridged) peptide in benign buffer (100 mM KCl, 50 mM PO4, pH 7.0) at the indicated temperature.

[c] The % α-helix at 25 °C was calculated based on an ellipticity value of -34,730 deg•cm^2•dmol^{-1} for a 100% helical 35-residue peptide derived from the equation $X_H^n = X_H^\infty(1 - k / n)$ where X_H^∞ is -39,500 deg•cm^2•dmol^{-1} for a helix of infinite length, n is the residue number per helix, and k is a wavelength dependent constant (2.5 at 222 nm) (31, 32).

[d] Mean residue molar ellipticities at 222 nm (deg•cm^2•dmol^{-1}). Measurements were performed on ~50 µM samples of oxidized peptide in benign buffer diluted 1:1 (v/v)with TFE.

[e] A ratio of $[\theta]_{222/208}$ greater than 1.0 is a measure of interacting α-helices in benign media, and a ratio of ~0.90 represents non-interacting α-helices in 50% TFE (34).

indicates that decreasing temperature stabilizes the folded structure for at least some of the analogs.

Under benign conditions and at 25 °C the ratio of ellipticities 222 nm/208 nm was greater than 1.0 for all of the peptides that exhibited molar ellipticities comparable to the theoretical maximum of -34,700 deg•cm^2•dmol^{-1} (31, 32) (Table VII-1), which is indicative of the presence of interacting α-helices (33, 34). The ratio of ellipticities 222 nm/208 nm at 25 °C for Arg19**a**, Gly19**a** and Orn19**a**, ranged between 0.96 to 1.00, which is consistent with the observation that for these peptides $P_f \neq 1$ under benign conditions at 25 °C. However, when 222 nm/208 nm was measured at 4.8 °C, the ratio was 1.03 or greater for Arg19**a**, Gly19**a**, Orn19**a** and Asp19**a**, indicating interacting α-helices at the lower temperature, which corroborates the observation that these peptides have the ability to completely populate the folded state. It was observed that the molar ellipticity in 50% TFE at 222 nm generally decreased by about 10% to 20%. This decrease in helicity in TFE may reflect the loss of synergism in the *n*-*π** transition (222 nm CD band) from interacting α-helices in the coiled-coil structure (34) since TFE has been shown to disrupt tertiary and quaternary structure, and to promote secondary structure (33-35). Interestingly, the ellipticity observed in 50% TFE at 25 °C decreased significantly for Glu19**a** from -35,000 deg•cm^2•dmol^{-1} in benign buffer to -25,300 deg•cm^2•dmol^{-1} in 50% TFE (Table VII-1) which suggests that charged residues are especially disruptive of α-helical structure in the form of monomeric α-helices in the non-polar environment of TFE, compared to disruption of α-helices involved in coiled-coil structures in benign medium which are stabilized by interhelical hydrophobic interactions. The ratio of ellipticities 222 nm/208 nm in 50% TFE decreased to 0.8 to 0.9 for all of the analogs (except Pro19**a**), which is characteristic of non-interacting α-helices. Pro had a devastating effect on coiled-coil structure (-1900 deg•cm^2•dmol^{-1}) compared to the monomeric α-helix in 50% TFE (-20,300 deg•cm^2•dmol^{-1}). This is in excellent agreement with previous studies on

---

Figure VII-3. Reversed-phase high performance liquid chromatographic separation of each analog to determine relative side-chain hydrophobicity at pH 7.0. The corresponding N-terminal *S*-carboxamidomethyl cysteine derivatives of the peptides were prepared to prevent oxidation, and chromatography was performed at 70 °C to prevent aggregation. Separation was carried out using a Zorbax 300 XDB C8 reversed-phase column as described in the methods section. Peptide A19**a** was included in each analog set as an internal reference standard, and the resulting retention times are shown in Table VII-2.

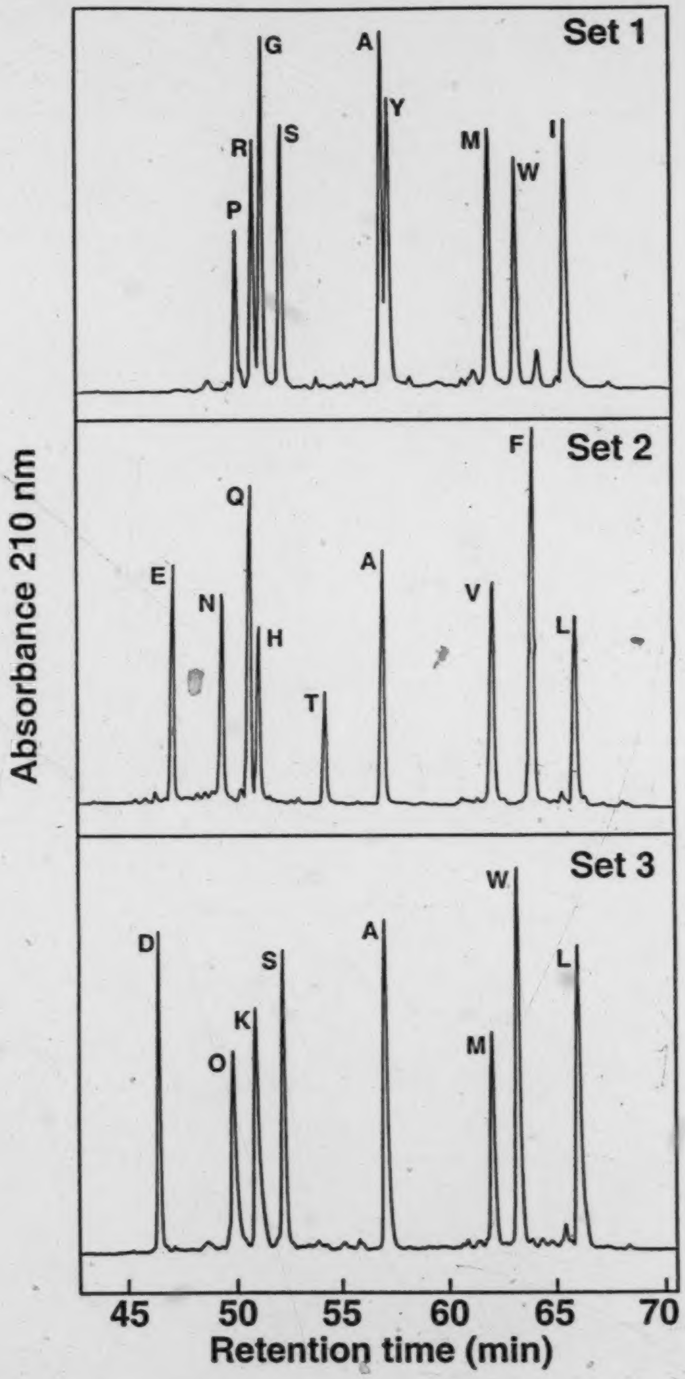

Set 1
P
R
G
S
A
Y
M
W
I
Set 2
E
N
Q
H
T
A
V
F
L
Set 3
D
O
K
S
A
M
W
L
Absorbance 210 nm
45
50
55
60
65
70
Retention time (min)

Table VII-2. Side-chain hydrophobicity data determined from amphipathic α-helical peptides, random coil peptides and N^{α}-acetyl amino acids.

Substituted Amino Acid[a]	Leu/Val face analogs (this study)			Ala face analogs	Random coil peptides	
	t_R[b] (min)	$\Delta t_R(\text{Gly})$[c] (min)	$\Delta t_R(\text{Gly})$ (normalized)[d]	$\Delta t_R(\text{Gly})$ (normalized)[e]	$\Delta t_R(\text{Gly})$ (normalized)[f]	ΔG(trans)[g] (kcal/mol)
Leu	64.6	13.1	100	100	100	2.32
Ile	64.3	12.8	98	101	92	2.46
Phe	62.8	11.3	86	103	100	2.44
Trp	62.2	10.7	82	100	105	3.07
Val	61.2	9.7	74	78	64	1.66
Met	61.1	9.6	73	76	67	1.68
Tyr	56.9	5.4	41	64	52	1.31
Ala	56.7	5.2	40	42	26	0.42
Thr	54.2	2.7	21	13	5	0.35
Ser	52.3	0.8	6	-5	-3	-0.05
Gly	51.5	0	0	0	0	0
His	51.4	-0.1	-1	8	26	0.18
Arg	51.2	-0.3	-2	-14	12	-1.37
Lys	51.2	-0.3	-2	-24	0	-1.35
Gln	51.0	-0.5	-4	-11	2	-0.30
Pro	50.5	-1.0	-8	-29	26	0.98
Orn	50.2	-1.3	-10	-	-	-
Asn	49.9	-1.6	-12	-29	-7	-0.82
Glu	47.8	-3.7	-28	-32	-12	-0.87
Asp	47.2	-4.3	-33	-57	-26	-1.05

Table VII-2 Legend

---

[a] Substituted amino acid at position 19**a** denotes the corresponding disulfide bridged coiled-coil. See Figure VII-1 for peptide amino acid sequence.

[b] t_R is the RP-HPLC retention time (pH 7.0 running buffer) of the X19**a** analogs (Figure VII-3), where the cysteine residue in each peptide was derivatized to *S*-carboxamidomethyl cysteine. See methods for separation conditions.

[c] Δt_R(Gly) is the difference in the RP-HPLC retention time of analog X19**a** relative to the Gly substituted analog. The non-polar face of this coiled-coil sequence consists of Val and Leu residues except at the substitution site.

[d] The normalized Δt_R(Gly) values from this study, where the Δt_R value for Leu19**a** = 100 and for Gly19**a** = 0.

[e] The normalized Δt_R(Gly) values of a substituted 18-residue monomeric amphipathic α-helix where the non-polar face consists of Ala residues except at the substitution site from Monera *et. al* (38) where the Δt_R value for the Leu substituted analog = 100 and that for the Gly substituted analog = 0.

[f] The normalized Δt_R(Gly) values of a substituted 8-residue random coil peptide from Guo *et al.* (40), where the Δt_R value for the Leu substituted analog = 100 and that for the Gly substituted analog = 0.

[g] ΔG(trans) is the free energy of transfer between octanol and water for the amino acid side-chain relative to glycine, from Eisenberg and McLachlan (41).

monomeric α-helices where essentially 0% α-helix was observed in benign buffer and 50% α-helix was observed in 50% TFE (36).

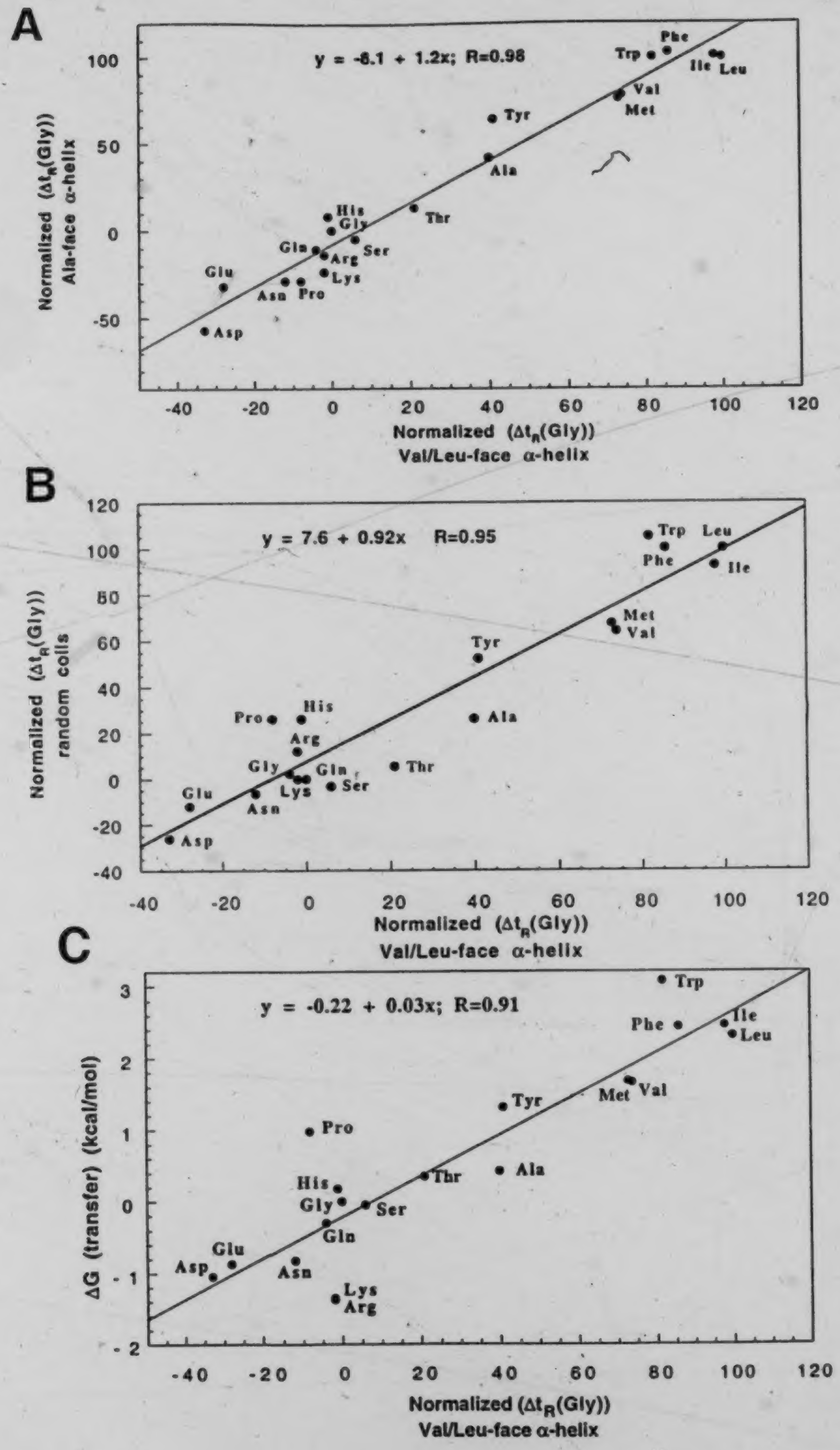
Side-chain hydrophobicity

Reversed-phase HPLC (RP-HPLC) was used to measure the hydrophobicity of the substituted amino acid side-chains since Van der Waals interaction of the amphipathic faces of the α-helices which form the coiled-coil in benign buffer is analogous to the interaction of the amphipathic face (preferred binding domain) of the α-helices with the reversed-phase C8 stationary phase (37). Moreover, while the relative hydrophobicity of the amino acid side-chains has been determined previously in other model systems (38-40), it was necessary to assess whether the relative hydrophobicity remained the same in the context of the amino acid sequence employed in the present study. The peptides were *S*-carboxamidomethylated to preclude re-oxidation of the cysteine residues during analysis, and RP-HPLC was performed at 70 °C to prevent aggregation of the Ile analog on the HPLC column observed at room temperature. The separation was also performed at 25 °C, the temperature at which the chemical denaturation experiments were performed, and the same elution order was observed except that the Ile substituted analog was not resolved. The RP-HPLC method used to determine side-chain hydrophobicity resulted in all of the analogs being well resolved under the elution conditions employed (Figure VII-3 and Table VII-2), with a 17.4 minute difference in retention times between the most hydrophilic analog (Asp19**a**, t_R = 47.2 min) and the most hydrophobic analog (Leu19**a**, t_R = 64.6 min). RP-HPLC separation of the analogs was performed individually as well as in 3 sets, with the common peptide A19**a** in each set as a retention time reference (Figure VII-3).

A summary of the side-chain hydrophobicity results are shown in Table VII-2. Comparison of these retention times with the relative order of hydrophobicity previously observed in a monomeric α-helix model system where the hydrophobic face consists of

---

Figure VII-4. Correlation of side-chain hydrophobicity data. *Panel A*: correlation of Δt_R(Gly) data obtained in a previous study by Monera et. al (38) using an Ala-face α-helical peptide with Δt_R(Gly) obtained in the present study using the Leu/Val-face peptide X19a. *Panel B*: correlation of Δt_R(Gly) data obtained in a previous study by Guo et. al (40) using a random coil peptide with Δt_R(Gly) from this study. *Panel C*: correlation of ΔG(transfer), the free energy of partitioning N-acetyl amino acid amides between a polar aqueous environment and octanol, from Eisenberg and McLachlan (41) and Δt_R(Gly) from this study.


A
y = -8.1 + 1.2x; R=0.98
Normalized (Δt_R(Gly)) Ala-face α-helix
Normalized (Δt_R(Gly)) Val/Leu-face α-helix
Phe
Trp
Ile
Leu
Val
Met
Tyr
Ala
Thr
His
Gly
Gln
Arg
Ser
Lys
Glu
Asn
Pro
Asp
100
50
0
-50
-40
-20
0
20
40
60
80
100
120
B
y = 7.6 + 0.92x R=0.95
Normalized (Δt_R(Gly)) random coils
Normalized (Δt_R(Gly)) Val/Leu-face α-helix
Trp
Leu
Phe
Ile
Met
Val
Tyr
Ala
Pro
His
Arg
Gly
Gln
Thr
Lys
Ser
Glu
Asn
Asp
120
100
80
60
40
20
0
-20
-40
-40
-20
0
20
40
60
80
100
120
C
y = -0.22 + 0.03x; R=0.91
ΔG (transfer) (kcal/mol)
Normalized (Δt_R(Gly)) Val/Leu-face α-helix
Trp
Phe
Ile
Leu
Tyr
Met
Val
Pro
Thr
Ala
His
Gly
Ser
Gln
Glu
Asp
Asn
Lys
Arg
3
2
1
0
-1
-2
-40
-20
0
20
40
60
80
100
120

Ala residues except at the substitution site (38) reveals an excellent correlation (R=0.98) of relative hydrophobicity (Figure VII-4, panel A). This indicates that the hydrophobicity of the amino acid side-chains substituted at position 19***a*** is model independent under reversed-phase HPLC conditions. Comparison of the retention times with those obtained from random coil peptides (Figure VII-4, panel B) (40) likewise revealed a very high correlation (R=0.95), indicating the hydrophobicity of a substituted side-chain measured when in an α-helix (Figure VII-4, panel A) is not perturbed by the structural context in which it occurs. Finally, the relative hydrophobicity values also correlated well (R=0.91) with ΔG(transfer) calculated by Eisenberg and McLachlan (41) from hydrophobic parameters of amino acids side chains measured by Fauchère and Pliska (39), where ΔG(transfer) is the free energy of transfer of the acetylated amino acid amide between an aqueous medium and octanol (Figure VII-4, panel C). Thus, these results taken together reveal that the differences in hydrophobicity between analogs is due solely to the change in the hydrophobicity of the substituted amino acid side chain. An exception is the analog Pro19**a**, which had a much lower retention time (Figure VII-4, panels B and C) than would be indicated by either the ΔG(transfer) value or from its hydrophobicity measured in random coil peptides (40), where the Pro and Ala substituted analogs had identical hydrophobicities. This can be understood since it is known that the carboxamido-methylated analogs partition as monomeric α-helices during RP-HPLC (42). The hydrophobic face of the amphipathic α-helix provides a preferred binding domain that is in van der Waals contact with the hydrophobic C8 matrix, which results in a retention time increase relative to a peptide of the same structure (α-helix) and amino acid composition but different amino acid sequence (37). The extremely helix-disrupting nature of the proline residue in fact introduces random coil character, thereby disrupting the preferred binding domain and resulting in a decrease in RP-HPLC relative retention time for this analog.

Protein stability

To determine the relative stability of the analogs, each peptide was denatured at pH 7.0 using both GdnHCl and urea denaturant in separate experiments, and the change in molar ellipticity at 222 nm was monitored using CD spectroscopy. The use of both GdnHCl and urea in the present study was especially useful since the specific contribution of the ionic interactions to protein stability is unmasked in urea (43, 44). Moreover, it has also been shown that using more than one method of denaturation can reveal the presence of potentially stable (populated) folding intermediates (45). It is important to note that it has previously been shown that the oxidized analogs that form three-stranded species under benign conditions (either exclusively or in conjunction with two-stranded species) undergo a transition to an exclusively two-stranded monomeric state prior to a loss of helicity as measured by CD spectroscopy (28). This characteristic of model peptide sequence X19**a** is crucial since it allows comparison of the effect on stability of side-chain substitution on a two-stranded oxidized monomer for all the analogs regardless of the oligomerization state observed in benign medium.

The urea and GdnHCl denaturation curves obtained are shown in Figure VII-5. The relative order of stability was generally the same in both denaturants, with the hydrophobic amino acid substituted analogs being the most stable, the charged residue substituted analogs the least stable and the polar side-chain substituted analogs being of intermediate stability (Table VII-3). The transitions were monophasic using either GdnHCl or urea, with single inflection points observed for all analogs, consistent with a monomolecular two-state unfolding process in both denaturants with intermediate states minimally populated. The transition midpoints were between 0.3 M and 3.3 M in GdnHCl, and between 0.6 M and 7.6 M in urea (Table VII-3). Comparison of the midpoints obtained in the two denaturants revealed an excellent correlation (R=0.99) between denaturant midpoints (Figure VII-6, panel A), with the more chaotropic reagent GdnHCl being about 2.3x more effective than urea as a denaturant.

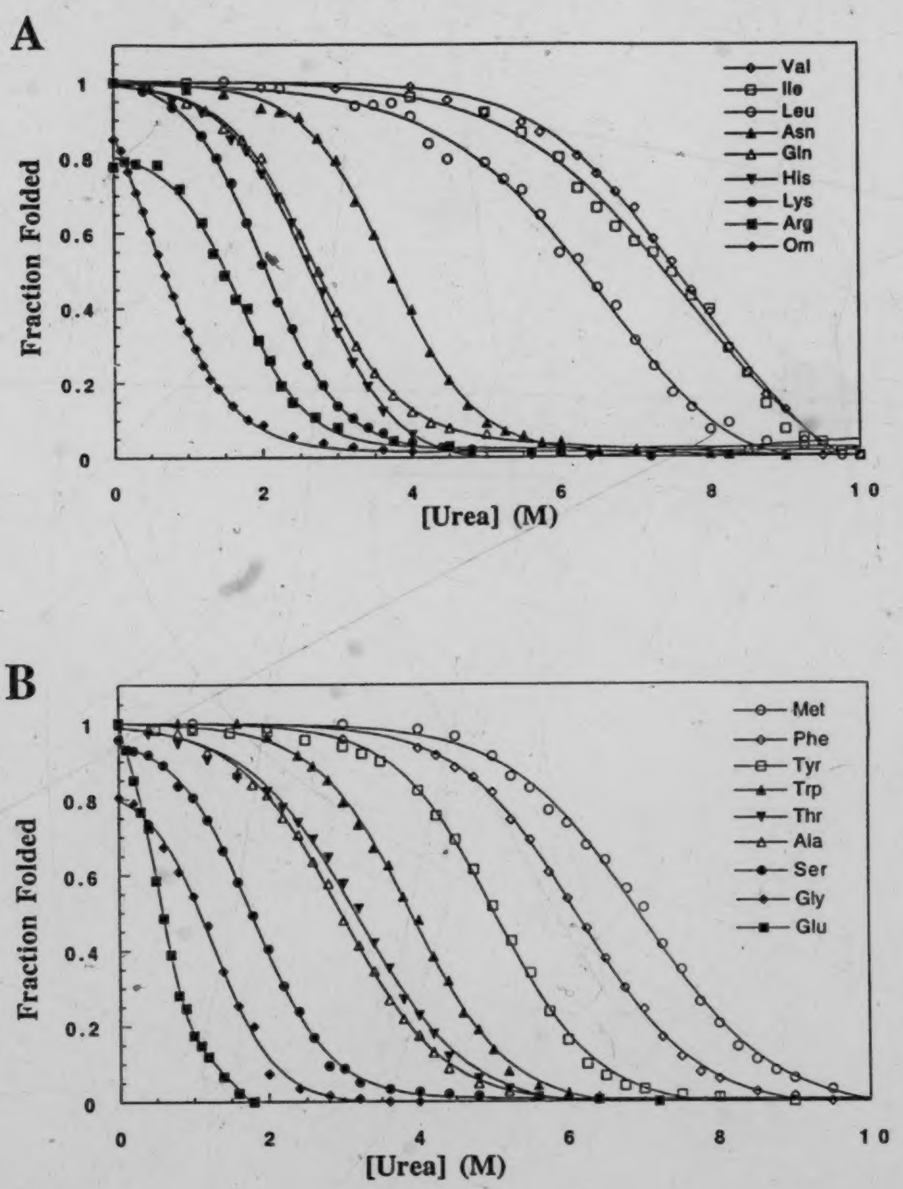
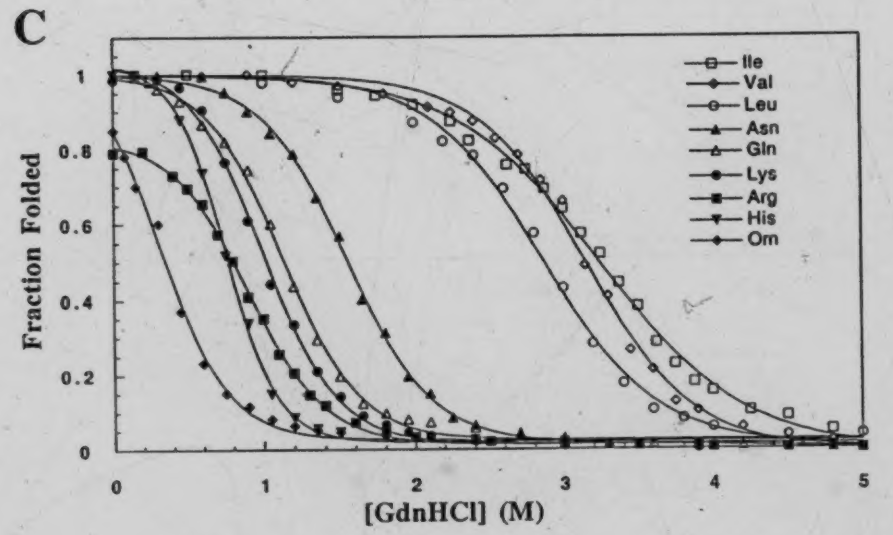


Figure VII-5. Chemical denaturation profiles. Panels A and B, urea denaturation profiles. Panels C and D, GdnHCl denaturation profiles. Panels A and C show the denaturation profiles for the Val, Ile, Leu, Asn, Gln, His, Lys, Arg and Orn analogs, and panels B and D show the denaturation profiles for the Met, Phe, Tyr, Trp, Thr, Ala, Ser, Gly and Glu analogs. A fraction folded value of 1.0 was used for all peptide analogs having negative molar ellipticities greater than 34, 300 in benign medium at 25 °C.


C
Fraction Folded
1
0.8
0.6
0.4
0.2
0
0
1
2
3
4
5
[GdnHCl] (M)
Ile
Val
Leu
Asn
Gln
Lys
Arg
His
Orn

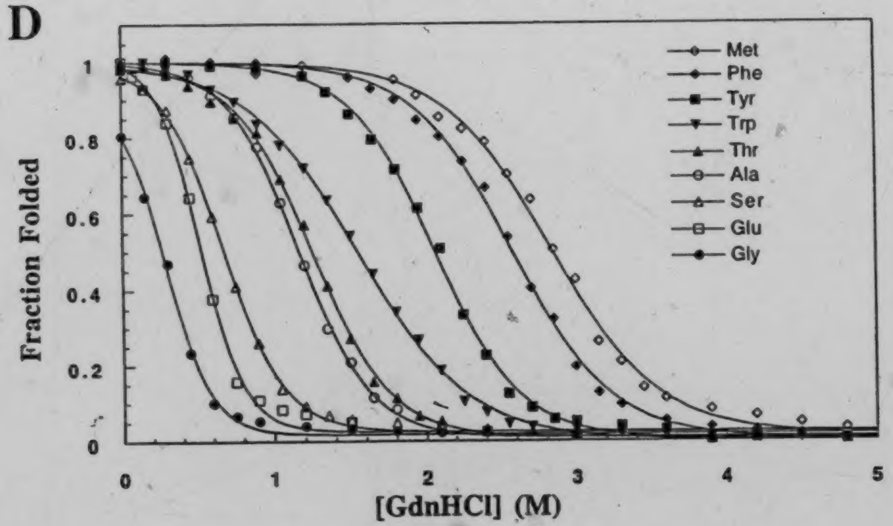

D
Fraction Folded
1
0.8
0.6
0.4
0.2
0
0
1
2
3
4
5
[GdnHCl] (M)
Met
Phe
Tyr
Trp
Thr
Ala
Ser
Glu
Gly

Table VII-3. Stability and oligomerization state data

Substituted Amino Acid[a]	Urea $[\text{Urea}]_{1/2}$ [b] (M)	Urea Slope [c] (m)	Urea ΔΔG [d] (Ala)	GdnHCl $[\text{GdnHCl}]_{1/2}$ [b] (M)	GdnHCl Slope [c] (m)	GdnHCl ΔΔG [d] (Ala)	Oligomer ratio [e]
Ile	7.40	0.58	3.29	3.26	1.18	3.88	4:6 (t:m)
Val	7.56	0.62	3.51	3.17	1.58	4.10	6:4 (t:m)
Met	6.94	0.77	3.31	2.85	1.53	3.43	9:1 (t:m)
Leu[f]	6.39	0.60	2.97	2.90	1.49	3.48	t
Phe	6.15	0.83	2.76	2.57	1.74	3.01	7:3 (t:m)
Tyr	5.02	0.93	1.89	2.07	2.09	2.08	t
Trp	3.91	1.01	0.90	1.54	1.65	0.78	m
Asn	3.73	0.96	0.71	1.54	1.94	0.85	m
Thr	3.18	0.86	0.18	1.25	2.13	0.20	8:2 (t:m)
Ala	2.97	0.91	0	1.17	2.53	0	4:6 (t:m)
Gln	2.74	0.94	-0.21	1.11	2.20	-0.13	t
His	2.64	1.11	-0.33	0.77	3.24	-1.15	t
Lys	2.05	1.13	-0.93	0.98	1.84	-0.40	m
Ser	1.80	1.11	-1.17	0.67	2.66	-1.28	3:7 (t:m)
Arg	1.51	1.00	-1.39	0.81	1.81	-0.77	m
Gly	1.08	1.01	-1.81	0.27	3.03	-2.49	4:6 (t:m)
Orn	0.68	1.31	-2.53	0.34	2.00	-1.88	m
Glu	0.60	2.17	-3.64	0.52	3.61	-1.98	1:1 (t:m)
Asp	-	-	-	-	-	-	-
Pro	-	-	-	-	-	-	-

Table VII-3 Legend

---

[a] Substituted amino acid at position 19**a** denotes the corresponding disulfide bridged coiled-coil. See Figure VII-1 for peptide amino acid sequence.

[b] $[\text{denaturant}]_{1/2}$ is the denaturation profile transition midpoint.

[c] m is the slope in the equation $\Delta G_u = \Delta G_{H2O} + m[\text{denaturant}]$ (74).

[d] $\Delta\Delta G_u(\text{Ala})$ is the free energy of unfolding of the analog relative to the Ala substituted analog, and is calculated from the equation $\Delta\Delta G_u(\text{Ala})=\{(m_{\text{Ala}}+ m_2)/2\}\{([\text{denaturant}]_{1/2})_{\text{peptide}} - ([\text{denaturant}]_{1/2})_{\text{Ala}}\}$ (75). Denaturation occured from the same two-stranded monomeric state for each analog (28), and the $\Delta\Delta G_u(\text{Ala})$ values reported are for two substitutions (one in each α-helix). Positive values indicate stabilities greater than that of the Ala substituted analog.

[e] Ratio of oligomerization states present by analyzing an aliquot of a 50 μM peptide in benign buffer by HPSEC as described in the methods. **m** denotes an exclusively monomeric oligomerization state, and **t** denotes an exclusively trimeric oligomerization state. All disulfide bridged peptides are monomeric prior to unfolding in denaturant (28).

[f] Denaturation of L19**a** was performed at peptide concentrations of 5.6 μM for the urea denaturation experiment and 2.5 μM for the GdnHCl denaturation experiment to ensure unfolding from a monomeric oligomerization state. Denaturation of all other analogs was performed at peptide concentrations of 50 μM.

The change in free energy of unfolding for each of the analogs relative to the Ala substituted peptide, ΔΔG(Ala), was measured using the GdnHCl and urea denaturation data as described in the methods section (Table VII-3). The importance of the hydrophobic core in modulating overall protein stability was reflected in the observation that a single amino acid mutation in peptide sequence X19**a** modulated the stability over a substantial 6 to 7 kcal/mol range (Table VII-3). The correlation of $\Delta\Delta G_{Ala}$ values obtained using the two denaturants was also excellent (R=0.97) (Figure VII-6, panel B). These results indicate that the relative stabilities of the analogs to chemical denaturation is nearly the same for each denaturant. However, the $\Delta\Delta G_{Ala}$ value obtained using urea as a denaturant for the charged side-chain analog Glu19**a** was significantly lower than that obtained from the corresponding GdnHCl denaturation experiment ($\Delta\Delta G_{Ala}$ = -3.64 kcal/mol in urea and -1.98 kcal/mol in GdnHCl; Table VII-3 and Figure VII-6, panel B). This observation is consistent with some shielding of ionic interactions of the negatively charged Glu carboxylate group at pH 7.0 due to ion-pairing with the hydrophobic guanidinium ion(44). As reported previously, the GdnHCl derived stabilities may provide a better value for the relative contribution of the hydrophobic interactions to coiled-coil stability, whereas urea denaturation studies allow the measurement of the net stability conferred by a given amino acid side-chain to overall protein stability (hydrophobic and charged interactions) (46). Thus, while comparable results were obtained using either GdnHCl or urea, the stability values obtained using urea provided the best measure of the effect of charged amino acid side-chain substitution on overall stability.

The effect on overall stability of mutating residues in the hydrophobic core of a coiled-coil or other protein is dependent on the difference in hydrophobicity of the amino acid side-chains, termed the hydrophobic effect, and the difference in the van der Waals interactions upon mutation, termed the packing effect(47). When the urea stability data ($\Delta\Delta G_u$ (Urea, Ala)) and the side-chain hydrophobicity data (Δt_r(Ala)) were compared, it was found that, in general, the stability of the X19***a*** analogs generally correlated with

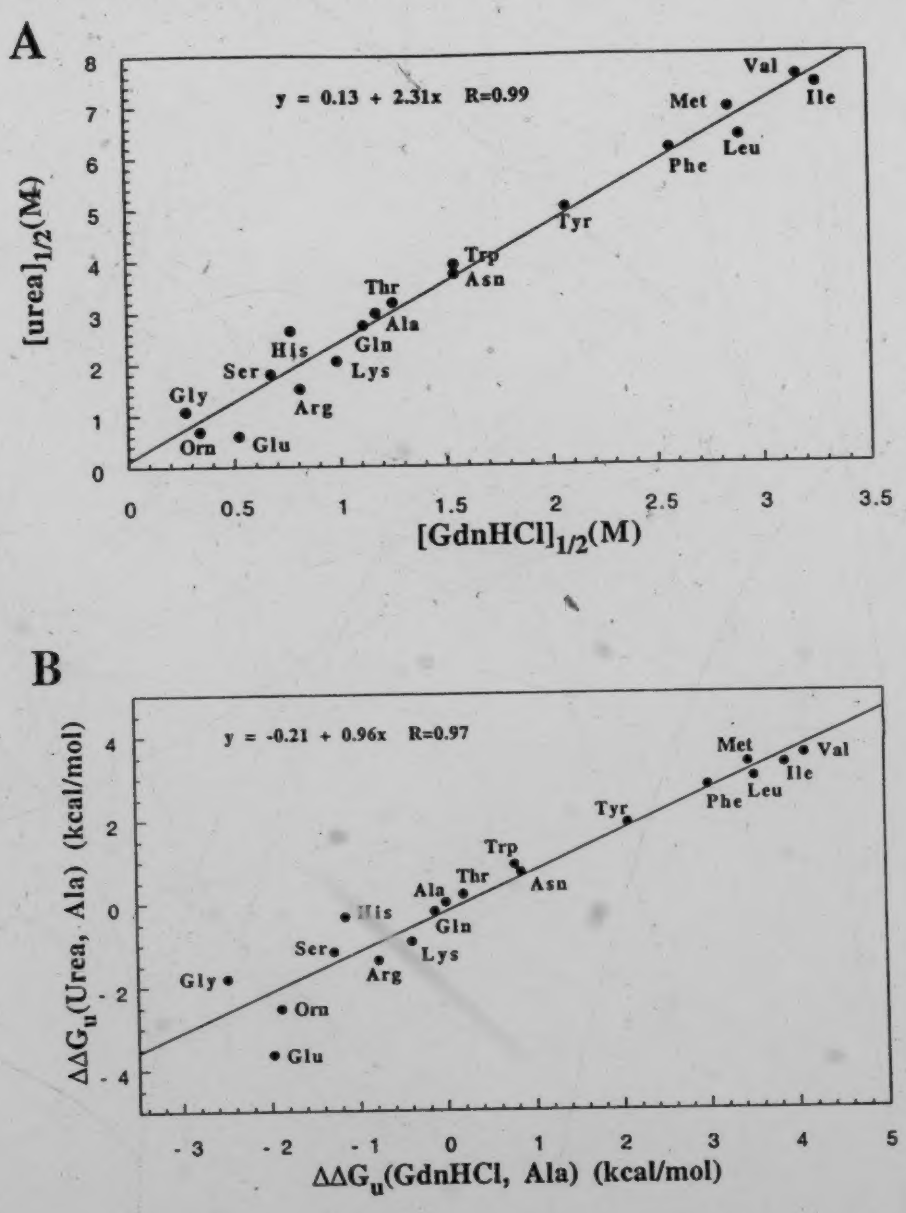


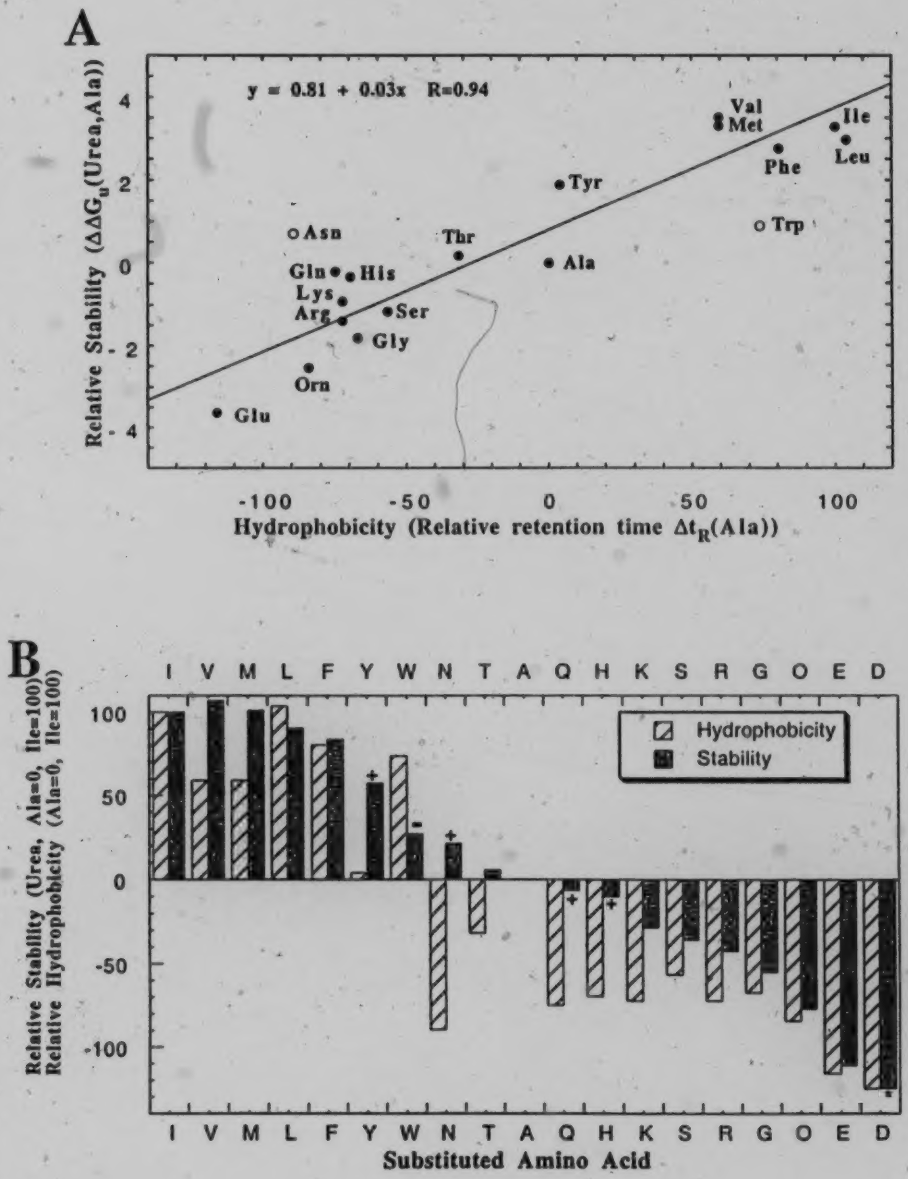
Figure VII-6. Comparison of chemical denaturation midpoints and relative stabilities obtained using GdnHCl and urea. Panel A: correlation of urea and GdnHCl chemical denaturation profile midpoints. Panel B: correlation of $\Delta\Delta G_u$(Ala) measured in GdnHCl and urea. The free energy of unfolding relative to the Ala substituted analog was measured using the equation: $\Delta\Delta G_u(\mathrm{Ala})=\{(m_{\mathrm{Ala}}+ m_2)/2\}\{([\mathrm{denaturant}]_{1/2})_2 - ([\mathrm{denaturant}]_{1/2})_{\mathrm{Ala}}\}$(77).

hydrophobicity of the substituted side-chain (R=0.94, Figure VII-7 panel A), with Asn19**a** and Trp19**a** being clear outliers (*vide infra*). Thus, 4 of the 6 most hydrophobic side-chains in the present study are also the four most stable X19**a** analogs, which in order of descending stability are V, I, M and L (Table VII-3). In a two-stranded coiled-coil the position "**a**" side-chains display packing geometries with the Cα-Cβ bond pointing out of the hydrophobic core, while the position "**d**" amino acid Cα-Cβ bond points directly at the adjacent α-helix (13). As a result β-branched amino acid side chains can be more stable in an "**a**" position than residues of similar or even greater hydrophobicity which are not β-branched, and also more stable than the same β-branched amino acid in a "**d**" position. We in fact observed that Ile19**a** was more stable than Leu19**a** due to this favorable packing effect, even though Leu is more hydrophobic (Table VII-2). These results are in agreement with previous studies in which it was shown that mutating Leu to Ile at position "**a**" in coiled-coils showed a significant packing effect in stabilizing coiled-coils, with $\Delta\Delta G_{Ala}$ values ranging from 0.59 to 1.03 kcal/mol (48). These results lend credence to the host/guest strategy employed since the model system was able to probe these differences in β-branched amino acid side-chain packing. Vinson and coworkers have recently reported a study designed to measure the energetic contribution to stability of 6 different amino acids in the hydrophobic core "**d**" position of a dimeric leucine zipper coiled-coil, and they found that the order of descending stability as determined by thermal denaturation experiments was L, M, I, V, C, A (49). Thus, their results demonstrated that Leu is the most stabilizing amino acid in the "**d**" position, in agreement with previous studies by Hodges and coworker (48), which is consistent with side-chain packing at a "**d**" position favoring Leu rather than β-branched amino acids in two-stranded coiled-coils.

It is generally accepted that the hydrophobic effect is the main factor in stabilizing the folded structure of globular proteins. Experiments designed to address this hypothesis have been performed in which Leu has been changed to Ala, with resultant decreases in stability that ranged from 1.7 kcal/mol to 6.2 kcal/mol (50). The lower value is comparable

---

Figure VII-7. Correlation of side-chain hydrophobicity and relative stability. *Panel A*: correlation of relative free energy of unfolding obtained from the urea denaturation experiments, $\Delta\Delta G_u$(Urea, Ala), with the RP-HPLC retention time of the analogs relative to the Ala substituted analog, Δt_R(Ala). The analogs used for linear extrapolation and calculation of the correlation coefficient are shown as closed circles, while the analogs substituted with Asn and Trp were outliers and were not used in the calculations, and are shown as open circles. *Panel B*: histogram representation of relative stability obtained from urea denaturation experiments, normalized such that the stability of Ile = 100 and Ala = 0, and the relative hydrophobicity obtained from RP- HPLC retention time data, normalized such that the hydrophobicity of Ile = 100 and Ala = 0. Analogs having a stability greater than indicated by apparent side-chain hydrophobicity are denoted "+" (Y,N,Q and H), while the analog (W) having a stability less than expected is denoted "-".


A
y = 0.81 + 0.03x R=0.94
Relative Stability (ΔΔG$_u$(Urea,Ala))
Hydrophobicity (Relative retention time Δt$_R$(Ala))
4
2
0
- 2
- 4
-100
-50
0
50
100
Val
Met
Ile
Leu
Phe
Tyr
Trp
Asn
Thr
Ala
Gln
His
Lys
Arg
Ser
Gly
Orn
Glu
B
Relative Stability (Urea, Ala=0, Ile=100)
Relative Hydrophobicity (Ala=0, Ile=100)
100
50
0
-50
-100
I V M L F Y W N T A Q H K S R G O E D
Hydrophobicity
Stability
Substituted Amino Acid

to the difference in the solvent transfer free energy ($\Delta\Delta G$(transfer)) of 1.9 Kcal/mol between the methyl and *t*-butyl side chains of Ala and Leu (Table VII-2), which is solely due to differences in hydrophobicity of the Leu and Ala side chains (39). The wide range between this lower value and the upper value has been ascribed to differences in the size of the cavity created when the Leu side chain is replaced with a methyl group, with the loss in stability correlating with the increase in cavity volume caused by the mutation (47). In the present study, the value of 1.6 kcal/mol/mutation observed for Leu→Ala substitution is comparable to $\Delta\Delta G$(transfer). This loss of stability is also comparable to that observed previously by Zhou *et al.*, (15), where mutation of two leucine residues at the corresponding position of a similar model coiled-coil protein resulted in a 1.8 kcal/mol/mutation loss in stability. This suggests that the differences in stability between the Leu and Ala analogs is mainly due to differences in side chain hydrophobicity, and that the two position "**a**" Ala side-chains can pack in a two-stranded coiled-coil without forming a cavity.

The Trp19**a** analog was observed to be ~2.3 kcal/mol less stable than would be predicted by relative hydrophobicity (Figure VII-7, panel A), indicated by the "-" sign in the histogram shown in Figure VII-7, panel B. Modeling shows this is due to incomplete burial in the hydrophobic core since only -50% of the indole ring appears to be buried in a two-stranded coiled-coil (Figure VII-9, panel C). Thus, the hydrophilic Nε1 of the indole side chain is buried in the core and the benzyl moiety, while contributing greatly to the hydrophobicity of the side-chain as measured by RP-HPLC, is not actually appreciably desolvated in the coiled-coil structure. In contrast, Tyr, Gln and His substitutions result in coiled-coils that are more stable (denoted by "+" signs in Figure VII-7, panel B) than would be expected based on side-chain hydrophobicity/hydrophilicity alone. Modeling of these residues in a two-stranded coiled-coil structure (modeling not shown) reveals that only the hydrophobic regions of these side-chains are buried, while the polar hydroxyl moiety of Tyr, imidazole nitrogens of His and Gln carboxamide group can remain at least partly solvated. However, the hydrophobicity measured by RP-HPLC for these side-chains entails desolvation of both the non-polar and polar regions of the Tyr, Gln and His side-chains in the reversed-phase matrix. Thus, side-chain burial in the isotropic reversed-phase matrix and in the coiled-coil hydrophobic core are not equivalent for the polar side-chains where selective desolvation of the more non-polar region of the side-chain can occur in the protein fold. The Asn substituted coiled-coil was also significantly more stable than expected (denoted by a "+" sign in Figure VII-8, panel B), and it known that the entire carboxamide group of Asn is buried in the hydrophobic core of similar coiled-coils (51). In this case the extra stabilization provided by Asn substitution in an **a** position is well documented as due to H-bonding of the Asn side-chain (13, 25, 51-53). In addition, there is evidence that the Asn side-chains can flip between two distinct H-bond conformations (24) which may provide some entropic compensation for burial of the polar carboxamide group (54).

Lysine at the "**a**" position of coiled-coils has been s[illegible] to occur in leucine zippers (55, 56). Modeling shows that the side-chain bearing the positively charged ε-amino group of Lys19**a** in a two-stranded coiled-coil is long enough to extend out of the hydrophobic core and allow at least partial solvation. The ornithine analog (Orn19**a**), the homologue of Lys19**a** in which the substituted side-chain bearing the positively charged amino group is shortened by one methylene group, served to probe the effect of bringing the charged amino group closer to the peptide backbone. As seen in Figure VII-7 panel A and in Table VII-3, Orn19**a** was 1.5 kcal less stable than Lys19**a** as measured in urea denaturation experiments. While Orn19**a** is slightly less hydrophobic than Lys19**a**, there is an additional decrease in stability observed above that expected due solely to decreased hydrophobicity. This observation is consistent with the δ-amino group of Orn being less solvated due to its closer proximity to the hydrophobic core.

While Glu was the most destabilizing amino acid substitution that still allowed coiled-coil formation under benign condition at room temperature, the acidic residues Glu and Asp do indeed occur in coiled-coils (20). Asp19**a** achieved a fully folded state only when the temperature was lowered to 4.8 °C (Table VII-1). However, destabilizing residues such as Asp and Glu can play an important role in controlling the preferential formation of heterodimeric species over homodimeric species *in vivo* since the presence of acidic side-chains at the dimerization interface is an effective mechanism to destabilize the incipient homodimer. For example, the basic region helix-loop-helix leucine zipper (b-HLH-LZ) oncoprotein c-Myc must preferentially heterodimerize with the b-HLH-LZ Max protein to bind DNA and activate transcription. Homodimerization of the c-Myc protein is prevented (and heterodimerization thus promoted by mass-action) by two Glu residues at "**a**" positions (57, 58), and in fact the leucine zipper domains of all the proteins known to date that interact with Max have a conserved acidic residue (either Glu or Asp) in the dimerization interface. In addition to disfavoring homodimer formation, these conserved acidic residue side-chains are also part of a molecular recognition mechanism that actively promotes specific heterodimer formation. Thus, recent NMR solution structure work by Lavigne *et al.* of the DNA binding c-Myc-Max heterodimer (59) showed conclusively that the previously postulated electrostatic interactions (58) between the two position "**a**" Glu side-chains in the c-Myc leucine zipper and a position "**d**" His side-chain in the Max leucine zipper are indeed responsible for the specificity of heterodimer formation. Glu is also seen in the hydrophobic core of kinesin molecules where it occurs in close proximity to Lys or Arg residues. The local environment is therefore more favorable than if substituted into the context of a strictly aliphatic hydrophobic core, and this context dependent modulation of stability has been suggested to be important in conferring upon the motor protein the ability to selectively carry out a movement signal (60).

Examination of the *m* values (Table VII-3), which are a measure of denaturant effectiveness for a given protein, revealed that the analogs fell into 3 groups; those containing very hydrophobic substitutions had the lowest *m* values, those with the most hydrophilic substitutions had the highest *m* values, and the remainder of the analogs displayed intermediate *m* values. The *m* values in GdnHCl were in general on the order of 2x those measured in urea. In a study in which denaturant *m* values from a large set of proteins was examined, it was found that *m* correlated (R=0.87 for GdnHCl) with the change in solvent accessible surface area (ASA, the amount of protein surface exposed to solvent upon unfolding) (61). The variation in *m* values can therefore be ascribed to differences in the unfolded states of the analogs, which is relevant because the denatured states of the analogs are equal in importance to the native states in determining protein stability (62). Thus, higher *m* values observed with the hydrophilic residues could be associated with unfolded states possessing relatively greater ASA and less residual structure, while the lower *m* values are observed with the hydrophobic side-chain substituted analogs are associated with unfolded states having relatively less ASA and more residual structure. In summary, it appears that the hydrophobic amino acid substitution may lead to increased residual structure in the unfolded state.

Protein stability versus natural occurrence at position "a"

Comparing the frequency of occurrence of amino acid residues at position "**a**" in naturally occurring coiled-coils(20) with their relative stability in model X19**a** from this study shows that the amino acid side-chains of the four most stable analogs (I19**a**, V19**a**, M19**a**, and L19**a**) also occur frequently at position "**a**" (Figure VII-8). Thus, the stability of naturally occurring coiled-coils is predominately due to the presence of highly hydrophobic side-chains in the hydrophobic core. Also, Gly and Glu, which give rise to the two least stable analogs, occur relatively infrequently at position "**a**" (Figure VII-8). In contrast to the highly hydrophobic amino acid side-chains, Figure 8 shows that there is no clear relationship between stability and frequency of occurrence for the side-chains of intermediate hydrophobicity. For example, Phe, Trp, Thr, Gln and His substitutions resulted in coiled-coils of appreciable stability, yet these residues are not used to a great

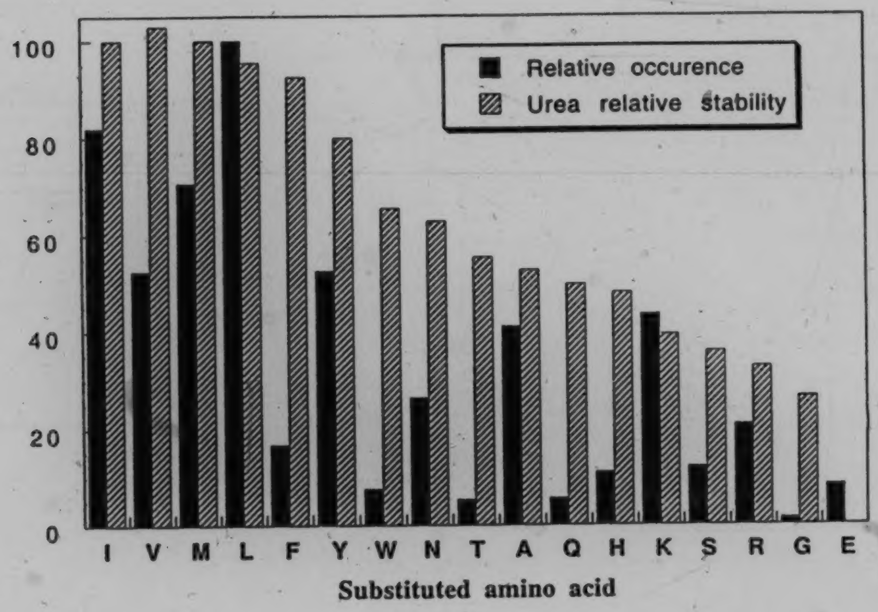


Figure VII-8. Comparison of side-chain occurence with relative side-chain stability. The relative stability obtained from urea denaturation experiments (Table VII-1) are normalized, with the stability of Ile = 100 and that of Glu = 0. The relative occurrence of the substituted amino acids in the hydrophobic core "a" position of coiled coils (Lupas et. al, 1991) are also normalized, with the occurrence of Leu = 100 and that of Glu = 0.

extent in naturally occurring coiled-coils. This may be partly due to some amino acids being not as prevalent in nature owing to the metabolic cost associated with their biosynthesis, and the use of some amino acids to control oligomerization state rather than stability (*vide infra*). Thus, we found that with the exception of Phe, nature has preferentially chosen hydrophobic amino acids for position "**a**" which provide maximal stability. The lack of a clear correlation between stability and frequency of occurrence for some of the analogs underscores the importance of the stability data obtained using our model peptide sequence in the present study for *de novo* design applications and for the creation of algorithms designed to predict the occurrence and stability of coiled-coils in naturally occurring peptides and proteins.

Oligomerization state

The structural polymorphism (oligomerization) of coiled-coil forming peptides has been documented for nearly six years (8, 56, 63). More relevant to this study are the examples of single amino acid replacement of the hydrophobic core position "**a**" wild-type Asn residue in the exclusively two-stranded GCN4 coiled-coil (8, 56, 63). To date the Ala, Abu, Gln, Nrl and Lys mutants have been characterized structurally (X-ray crystal structure) and hydrodynamically (oligomerization state). Only the wild type Asn and the Lys mutant are known to populate exclusively the two-stranded state in that system (56), while all the other mutants are known to populate both the two- and three-stranded states (8, 56, 63).

The oligomeric state(s) populated by the different analogs from the present study under benign conditions are presented in Table VII-3. The oligomerization state of the analogs under benign conditions was determined by high performance size-exclusion chromatography (HPSEC), and the molecular weights of the HPSEC peaks assigned as monomers (two-stranded) or trimers (three-stranded) was confirmed by sedimentation equilibrium ultracentrifugation. The quaternary structure of the oligomers was previously assessed in sedimentation velocity experiments that showed the oxidized peptides employed throughout this study form two-stranded monomers (Figure VII-2a) and elongated three-stranded trimers (Figure VII-2b) (28). We found that amino acid substitution at a single hydrophobic core "**a**" position in the 5-heptad model coiled-coil could switch the oligomerization state from exclusively monomeric to exclusively trimeric, with 9 of the 19 hydrophobic core position **a** substitutions directing the oligomerization state to a unique two- or three-stranded conformation (Table VII-3). This clearly demonstrates the importance of the effect of packing in the hydrophobic core structure on oligomerization state. Thus, at protein concentrations equivalent to those under which the chemical denaturation experiments were carried out (50 μM) and under benign conditions, Leu19**a**, Tyr19**a**, Gln19**a** and His19**a** were found to be exclusively three-stranded (trimeric), while Asn19**a**, Lys19**a**, Orn19**a**, Arg19**a**, and Trp19**a** were exclusively two-stranded (monomeric). These results are especially relevant for our understanding of the polymorphism of coiled-coils and its prediction since it increases the database to all proteogenic amino acids except Cys at position "**a**" in a model peptide.

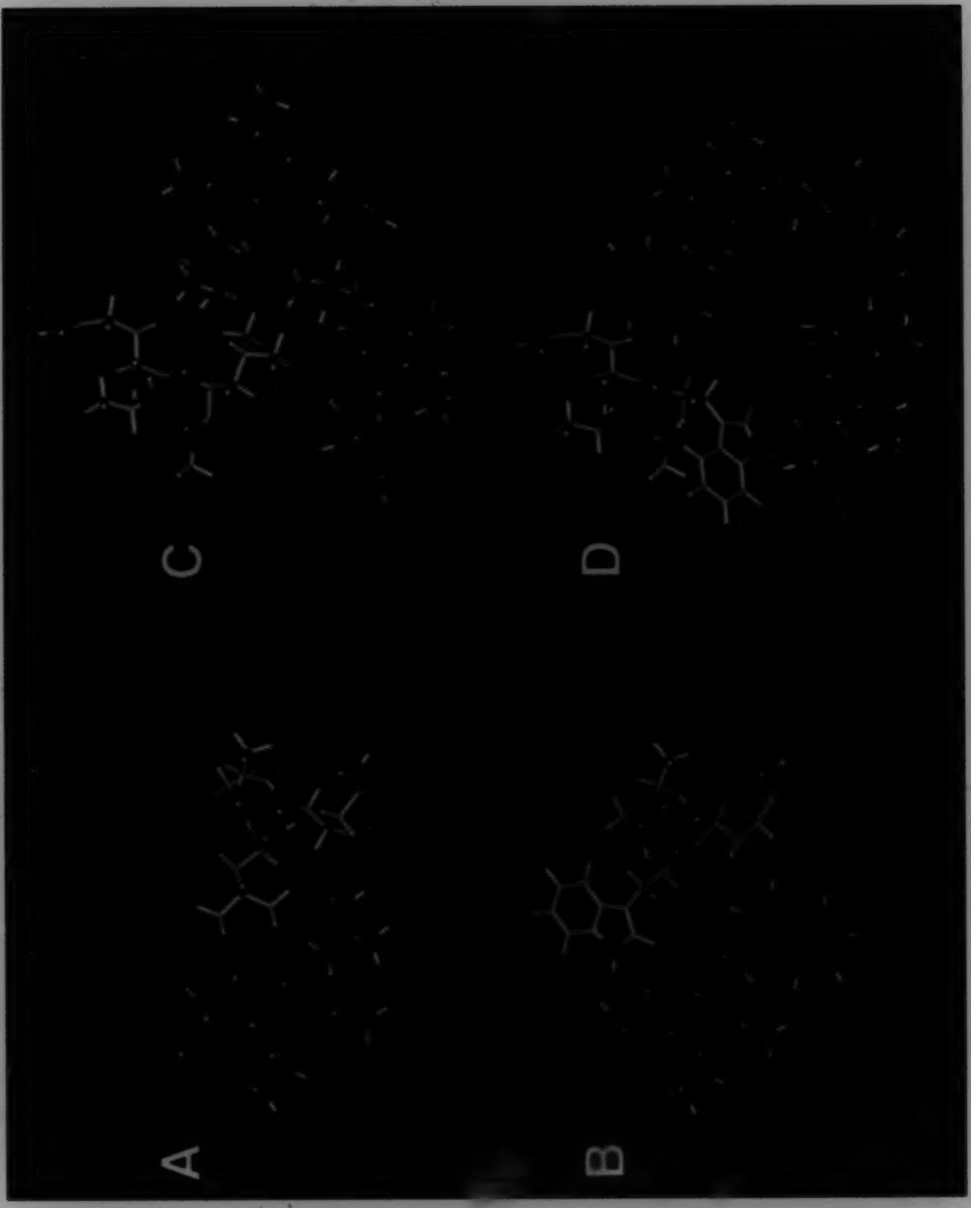
It should be noted that the observed oligomerization states in this study are likely somewhat context dependent, however, since positions "**e**" and "**g**" are also involved in the hydrophobic interface and have been shown to play a role in determining oligomerization state (64-68). On the other hand, it is of particular interest to note that we also observe two-stranded specificity for the Asn and Lys analogs. This structural specificity imparted by Asn and Lys has been ascribed to the destabilization of the three-stranded state by unoptimized buried H-bonds (Asn) and unfavorable desolvation of the charged Nε of Lys (56). These unfavorable events are relieved in the two-stranded state where the Asn can form a interhelical side-chain H-bond and the Nε of Lys is more favorably solvated (56). This selective destabilization of one state can be seen as a general mechanism for the structural specificity imparted by some of the side-chains. Hence, it could be predicted (and as proposed by others (56)) that the Arg analog should show the same specificity for a two-stranded state, and in fact this is observed here for both the Arg and Orn analogs. While it is tempting to suggest that all the analogs that are potentially charged at pH 7.0 should show specificity for a two-stranded structure, it is interesting to note the Glu analog is an exception in that both the two- and three-stranded states are populated. However, the pKa of Glu side-chain carboxylate group at this position may be perturbed, and must be determined before an explanation for this observation can be proposed.

Our current efforts on the modeling of the Leu analog in both two- and three-stranded oligomerization states (Figure VII-9, panels A and B) indicate that the Leu side-chains pack asymmetrically in the two-stranded state ($\chi1$ = gauche + , $\chi2$ = trans and $\chi1$ = trans, $\chi2$ = gauche -) in order to satisfy good van der Waals interaction (avoid the formation of a cavity). In the three-stranded state it is also possible to avoid cavity formation and optimize the packing of side-chains which adopt an asymmetric arrangement (2 side-chains with $\chi1$ = gauche + , $\chi2$ = trans and one with $\chi1$ = trans, $\chi2$ = gauche -). Interestingly, the Leu side-chains in the two-stranded fold are on average 82 % buried while they are 96 % buried in the three-stranded fold, suggesting in part why this analog selectively folds as a three-stranded coiled-coil. The calculation of the accessible surface area of Ile side-chains at position "**a**" in two- and three-stranded states reveals similar extents of side-chain burial as observed for the Leu analog. Nevertheless, the Ile analog is observed to populate both the two- and three-stranded states. This can be understood, however, since burial in a three-stranded structure of hydrophobic surface area is done with a conformational entropy penalty for an Ile side-chain that is significantly larger than that incurred for a Leu side-chain (69). This should destabilize the three-stranded state of the Ile analog (making it relatively less populated) to the advantage of the two-stranded state.

Side-chain packing in relation to cavity formation is also an important consideration in determining the oligomerization state. Previously, Hodges and coworkers showed that the oligomerization state of coiled-coils can be modulated by the relative positioning of Ala residues in the hydrophobic core(70). They found that the small Ala side-chain methyl group controls the formation of two-stranded or four-stranded α-helical coiled-coils by creating a large cavity if four Ala residues are on the same plane, which disfavors tetramer formation and results in a two-stranded coiled-coil. Gonzalez and coworkers (56) used the principle of cavity formation leading to lowered stability to engineer an allosteric switch in GCN-4 (71). They changed the position 16**a** Asn residue to Ala, and the resulting coiled-coil was found to populate both two-and three-stranded oligomerization states. A nearly exclusively three-stranded structure was formed upon addition of benzene or cyclohexane, which were shown in high resolution structures to fill a cavity created in the 3-stranded structure by substitution of the smaller Ala side-chain. Similarly, the packing effect of placing the large indole ring of Trp in the hydrophobic core is also important in controlling resultant oligomerization state. Modeling of Trp19**a** in both two- and three-stranded conformations (Figure VII-9, panels C and D) shows that packing of the bulky indole ring side-chain in a three-stranded conformation is highly unfavorable due to the formation of a large cavity (114 Å^3), and this is likely the driving force causing the formation of exclusively two-stranded molecules for this analog. In agreement with this proposal is the observation that in the synthetic peptide coil-Ser, the anti-parallel orientation of one of the 3 helices in the bundle was due to the inability to pack three Trp residue side-chains in the same core layer (72). However, while Trp does not pack well, and is in fact somewhat destabilizing relative to its hydrophobicity as discussed earlier, Trp does in fact occur in the hydrophobic core of natural coiled-coils (20), and has been shown to affect, for example, inter-helix spacing in gp41 trimers (73). Thus, the presence of the large Trp side-chain in the core contributes to helix orientation, protein function, as well as to oligomerization state.

Mixtures of two- and three-stranded coiled-coils were observed for 9 of the analogs under benign conditions (Table VII-3), and the observation of multiple

Figure VII-9. Modeling of the Leu and Trp analogs as two- and three-stranded coiled-coils. Panels A and B, modeling of the Leu analog. Panels C and D, modeling of the Trp analog.


A
B
C
D

conformations from a single sequence implies that the two- and three-stranded structures have comparable stabilities (74). The mixture of two- and three-stranded oligomerization states observed for the Asn16Ala GCN4 mutant of Gonzalez and coworkers mentioned previously (71) was also observed for the Ala analog in this study (Table VII-3). It has been reported that a GCN4 leucine zipper mutant in which Asn16 was changed to aminobutyric acid (Abu) could exist in either both two- or three-stranded coiled-coil conformations, and high resolution structures showed that the mutant was accommodated in the two distinct structures by utilizing different residues in each conformation to form similarly-shaped packing surfaces (63). It seems likely that a similar phenomenon is responsible for the polymorphism observed in the present study.

Distinctive residue propensities (frequencies of occurrence) for oligomerization states agreed with our experimentally measured effects of amino acid substitution on our model coiled-coil protein for only some of the analogs which exhibited unique oligomerization states. Thus, Asn, Lys and Arg occur 2.4x, 6.9x and 12.9x, respectively, more often at the "**a**" position of two-stranded coiled-coils than three-stranded coiled-coils, while Gln occurs 11.8 times more in three-stranded coiled-coils than in two-stranded coiled-coils(22). However, for the Leu19**a**, Tyr19**a** and His19**a** analogs there was a nearly equal occurrence of the residues in two- and three-stranded coiled-coils, while Trp occurred only once in the database, and was actually found in a parallel three-stranded molecule. Thus, as shown earlier in relation to overall coiled-coil stability, differences in the statistical occurrence of a given residue at position **a** of two- and three-stranded coiled-coils does not necessarily predict the actual oligomerization state.

C. Conclusion

These findings are the first comprehensive quantitative assessment of the effect of side-chain substitution in the hydrophobic core on the stability of two-stranded coiled-coils, and also provide information on the effect of side-chain hydrophobicity and packing in the hydrophobic core on the oligomerization state. This information will be invaluable for the development of coiled-coil predictive algorithms, and may also find use in estimating the stability of heterodimers. The effects on stability and oligomerization state of amino acid substitution observed in our model coiled-coil in general corroborate results previously made using natural protein sequences, which shows that the data obtained using this host/guest system will be generally applicable. Side-chain packing as a determinant of the relative stability of alternative oligomerization states appears to be crucial, as illustrated by the molecular modeling of the Leu and Trp substituted analogs in the present study. Efforts to model all 20 analogs in both two- and three-stranded oligomerization states are in progress, and these results will be presented in detail elsewhere.

D. References

1. Crick, F. H. C. (1953) *Acta Crystallogr.* **6**, 689-698.
2. Kohn, W. D., Mant, C. T., and Hodges, R. S. (1997) *J. Biol. Chem.* **272**, 2583-2586.
3. Kohn, W. D. and Hodges, R. S. (1998) *Trends Biotech.* **16**, 379-389.
4. Adamson, J. G., Zhou, N. E., and Hodges, R. S. (1993) *Curr. Opin. Biotech.* **4**, 428-437.
5. Kammerer, R. A. (1997) *Matrix Biol.* **15**, 555-565.
6. Cohen, C. and Parry, D. A. D. (1990) *Proteins: Struct. Funct. Genet.* **7**, 1-15.
7. Alber, T. (1992) *Curr. Opin. Genet. Dev.* **2**, 205-210.
8. Hodges, R. S. (1996) *Biochem. Cell Biol.* **74**, 133-154.
9. Hodges, R. S. (1992) *Current Biology.* **2**, 122-124.
10. Baxevanis, A. D. and Vinson, C. R. (1993) *Curr. Opin. Genet. Dev.* **3**, 278-285.

11. Betz, S. F., Bryson, J. W., and DeGrado, W. F. (1995) *Curr. Opin. Struct. Biol.* **5**, 457-463.

12. Lupas, A. (1996) *Trends Biochem. Res.* **21**, 375-382.

13. Harbury, P. B., Zhang, T., Kim, P. S., and Alber, T. (1993) *Science.* **262**, 1401-1407.

14. Hodges, R. S., Zhou, N. E., Kay, C. M., and Semchuk, P. D. (1990) *Pept. Res.* **3**, 123-137.

15. Zhou, N. E., Kay, C. M., and Hodges, R. S. (1992) *Biochemistry.* **31**, 5739-5746.

16. Zhou, N. E., Kay, C. M., and Hodges, R. S. (1992) *J. Biol. Chem.* **267**, 2664-2670.

17. Thompson, K. S., Vinson, C. R., and Freire, E. (1993) *Biochemistry.* **32**, 5491-5496.

18. Greenfield, N. J. and Hitchcock-DeGregori, S. E. (1995) *Biochemistry.* **34**, 16797-16805.

19. Lupas, A. (1997) *Curr. Opinion Struct. Biol.* **7**, 388-393.

20. Lupas, A., Van Dyke, M., and Stock, J. (1991) *Science.* **252**, 1162-1164.

21. Berger, B., Wilson, D. B., Wolf, E., Tonchev, T., Milla, M., and Kim, P. S. (1995) *Proc. Natl. Acad. Sci.* **92**, 8259-8263.

22. Woolfson, D. N. and Alber, T. (1995) *Protein Sci.* **4**, 1596-1607.

23. Wolf, E., Kim, P. S., and Berger, B. (1997) *Protein Science.* **6**, 1179-1189.

24. Junius, F. K., Mackay, J. P., Bubb, W. A., Jensen, S. A., Weiss, A. S., and King, G. F. (1995) *Biochemistry.* **34**, 6164-6174.

25. Lumb, K. J. and Kim, P. S. (1995) *Biochemistry.* **34**, 8642-8648.

26. Hodges, R. S., Saund, A. K., Chong, P. C. S., St.-Pierre, S. A., and Reid, R. E. (1981) *J. Biol. Chem.* **256**, 1214-1224.

27. Chao, H., Bautista, D. L., Litowski, J., Irvin, R. T., and Hodges, R. S. (1998) *J. Chrom. B.* **715**, 307-329.

28. Wagschal, K., Tripet, B., and Hodges, R. S. (1999) *J. Mol. Biol.* **285**, 785-803.

29. Monera, O. D., Kay, C. M., and Hodges, R. S. (1994) *Biochemistry.* **33**, 38692-3871.

30. O'Shea, E. K., Rutkowski, R., and Kim, P. S. (1989) *Science.* **243**, 538-542.

31. Chang, C. T., Wu, C.-S. C., and Yang, J. T. (1978) *Anal. Biochem.* **91**, 13-31.

32. Chen, Y.-H., Yang, J. T., and Chau, K. H. (1974) *Biochemistry.* **13**, 3350-3359.

33. Cooper, T. M. and Woody, R. W. (1990) *Biopolymers.* **30**, 657-676.

34. Lau, S. Y. M., Taneja, A. K., and Hodges, R. S. (1984) *J. Biol. Chem.* **259**, 13253-13261.

35. Sönnichsen, F. D., Van Eyk, J. E., Hodges, R. S., and Sykes, B. D. (1992) *Biochemistry.* **31**, 8790-8798.

36. Zhou, N. E., Monera, O. D., Kay, C. M., and Hodges, R. S. (1994) *Protein Peptide Lett.* **1**, 114-119.

37. Zhou, N. E., Mant, C. T., and Hodges, R. S. (1990) *Peptide Research.* **3**, 8-20.

38. Monera, O. D., Sereda, T. J., Zhou, N. E., Kay, C. M., and Hodges, R. S. (1995) *J. Peptide Sci.* **1**, 319-329.

39. Fauchère, J.-L. and Pliska, V. (1983) *Eur. J. Med. Chem.* **18**, 369-375.

40. Guo, D., Mant, C. T., Taneja, A. K., Parker, R., and Hodges, R. S. (1986) *J. Chromatogr.* **359**, 499-517.

41. Eisenberg, D. and McLachlan, A. D. (1986) *Nature.* **319**, 199-203.

42. Blondelle, S. E., Ostresh, J. M., Houghton, R. A., and Pérez-Payá, E. (1995) *Biophys. J.* **68**, 351-359.

43. Monera, O. D., Zhou, N. E., Kay, C. M., and Hodges, R. S. (1993) *J. Biol. Chem.* **268**, 19218-19227.

44. Kohn, W. D., Kay, C. M., and Hodges, R. S. (1995) *Prot. Sci.* **4**, 237-250.

45. Morjana, N. A., McKeone, B. J., and Gilbert, H. F. (1993) *Proc. Natl. Acad. Sci.* **90**, 2107-2111.

46. Monera, O. D., Kay, C. M., and Hodges, R. S. (1994) *Protein Sci.* **3**, 1984-1991.

47. Eriksson, A. E., Baase, W. A., Zhang, X.-J., Heinz, D. W., Blaber, M., Baldwin, E. P., and Matthews, B. W. (1992) *Science.* **255**, 178-183.

48. Zhu, B.-Y., Zhou, N. E., Kay, C. M., and Hodges, R. S. (1993) *Protein Sci.* **2**, 383-394.

49. Moitra, J., Szilák, L., Krylov, D., and Vinson, C. (1997) *Biochemistry.* **36**, 12567-12573.

50. Matthews, B. W. (1993) *Structural and genetic analysis of protein stability*, in *Annu. Rev. Biochem.*, C.C. Richardson, Editor, p. 139-160.

51. O'Shea, E. K., Klemm, J. D., Kim, P. S., and Alber, T. (1991) *Science.* **254**, 539-544.

52. Greenfield, N. J., Montelione, G. T., Farid, R. S., and Hitchcock-DeGregori, S. E. (1998) *Biochemistry.* **37**, 7834-7843.

53. Lavigne, P., Crump, M. P., Cagné, S. M., Sykes, B. D., Hodges, R. S., and Kay, C. M. (1997) *Tech. in Prot. Chem.* **8**, 617-624.

54. Mackay, J. P., Shaw, G. L., and King, G. F. (1996) *Biochemistry.* **35**, 4867-4877.

55. Glover, J. N. M. and Harrison, S. C. (1995) *Nature.* **373**, 257-261.

56. Gonzalez, L., Jr., Woolfson, D. N., and Alber, T. (1996) *Nature Struct. Biol.* **3**, 1011-1018.

57. Muhle-Goll, C., Gibson, T., Schuck, P., Schubert, D., Nalis, D., Nilges, M., and Pastore, A. (1994) *Biochemistry.* **33**, 11296-11306.

58. Lavigne, P., Kondejewski, L. H., Michael E. Houston, J., Sonnichsen, F. D., Lix, B., Sykes, B. D., Hodges, R. S., and Kay, C. M. (1995) *J. Mol. Biol.* **254**, 505-520.

59. Lavigne, P., Crump, M. P., Cagné, S. M., Hodges, R. S., Kay, C. M., and Sykes, B. D. (1998) *J. Mol. Biol.* **281**, 165-181.

60. Tripet, B., Vale, R. D., and Hodges, R. S. (1997) *J. Biol. Chem.* **272**, 8946-8956.

61. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) *Protein Sci.* **4**, 2138-2148.

62. Muñoz, V. and Serrano, L. (1996) *Folding and Design.* **1**, R71-R77.

63. Gonzalez, L., Jr., Brown, R. A., Richardson, D., and Alber, T. (1996) *Nature Struct. Biol.* **3**, 1002-1009.

64. Alberti, S., Oehler, S., von Wilcken-Bergmann, B., and Müller-Hill, B. (1993) *EMBO J.* **12**, 3227-3236.

65. Beck, K., Gambee, J. E., Kamawal, A., and Bächinger, H. P. (1997) *EMBO J.* **16**, 3767-3777.

66. Kammerer, R. A., Schulthess, T., Landwehr, R., Lustig, A., Fischer, D., and Engel, J. (1998) *J. Biol. Chem.* **273**, 10602-10608.

67. Kohn, W. D., Kay, C. M., and Hodges, R. S. (1998) *J. Mol. Biol.* **283**, 993-1012.

68. Zeng, X., Zhu, H., Lashuel, H. A., and Hu, J. C. (1997) *Prot. Sci.* **6**, 2218-2226.

69. Lee, K. H., Xie, D., Freire, E., and Amzel, M. (1994) *Proteins: Struct. Funct. Genet.* **20**, 68-84.

70. Monera, O. D., Sönnichsen, F. D., Hicks, L., Kay, C. M., and Hodges, R. S. (1996) *Protein Eng.* **9**, 353-363.

71. Gonzalez, L., Jr., Plecs, J. J., and Alber, T. (1996) *Nature Struct. Biol.* **3**, 510-515.

72. Lovejoy, B., Choe, S., Cascio, D., McRorie, D. K., DeGrado, W. F., and Eisenberg, D. (1993) *Science.* **259**, 1288-1293.

73. Weissenhorn, W., Dessen, A., Harrison, S. C., Skehel, J. J., and Wiley, D. C. (1997) *Nature.* **387**, 426-430.

74. Dill, K. A., Bromberg, S., Yue, K., Fiebig, K. M., Yee, D. P., Thomas, P. D., and Chan, H. S. (1995) *Protein Sci.* **4**, 561-602.

75. Chong, P. C. S. and Hodges, R. S. (1982) *J. Biol. Chem.* **257**, 2549-2555.

76. Pace, C. N. (1986) *Methods Enzymol.* **131**, 266-280.

77. Sali, D., Baycroft, M., and Fersht, A. R. (1991) *J. Mol. Biol.* **220**, 779-788.

78. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins: Struct. Funct. Genet.* **11**, 281-296.

Chapter VIII

The Effect of Side-Chain Characteristics on the Stability and Oligomerization State of a *De Novo* Designed Model Coiled-Coil: Twenty amino acid substitutions in a "d" position.

A version of this chapter is presently being revised for publication.

A. Introduction

The α-helical coiled-coil structural motif is characterized by a regularly repeating heptad unit typically designated **(abcdefg)**$_n$ (McLachlan and Stewart, 1975), in which positions **a** and **d** are occupied primarily by hydrophobic residues, positions **e** and **g** by charged residues and positions **b, c** and **f** by polar and charged residues. When configured in an α-helix, the periodic 3-4 repeat (Hodges *et al.*, 1972) of hydrophobic residues occurring in the **a** and **d** positions creates a nonpolar strip along the helix which can pack together with one or more other similar strands via the classical "knobs into holes" packing arrangement to bury their hydrophobic surface areas and stabilize the α-helical fold (Crick, 1953). This packing arrangement also deforms the individual α-helices slightly from their normal α-helical geometry of 3.6 residues/turn to 3.5 residues/turn and thus create a slight left-hand superhelical twist or the rope-like topology often observed in the crystal structures of two-, three- and four-stranded coiled-coils.

Extensive structural studies using *de novo* designed model coiled-coil peptides as well native coiled-coil sequences have clearly demonstrated the importance of the hydrophobic core **a** and **d** residues for the folding and stability of the coiled-coil structure (Hodges *et al.*, 1981, 1990; Hu *et al.*, 1990; Lau *et al.*, 1984; Moitra *et al.*, 1996; Zhou *et al.*, *1992a,b*; Zhu *et al.*, 1992, 1993). Interestingly, these studies also revealed that not only is the type of residue important (i.e. hydrophobic vs. polar) but also its placement in either an **a** or **d** position. For example, it was observed that the long chain aliphatic amino acid residues such as leucine provided significantly greater stability than that of the similarly sized hydrophobic residues of Ile and Val at a **d** position (Zhou *et al.*, 1992a,b; Zhu *et al.*, 1993); whereas the β-branched hydrophobic residues of Ile and Val appear to be more stable at an **a** position (Zhu *et al.*, 1993). These results are in good agreement with the observed statistical frequency difference between these two positions (Cohen and Parry, 1990). The solution to this positional preference enigma was solved in 1991 with the first high resolution X-ray crystal structure of the coiled-coil dimerization domain of the transcription factor GCN4 protein (O'Shea *et al.*, 1991). Because of the parallel and in-register nature of the α-helices, and their "knobs into holes" packing arrangement, the **a** [illegible] **d** positions are not similar (as originally believed), but rather structurally distinct. That [illegible] the **a** position shows a structural geometry in which the C_α-C_β vector extends parallel to the C_α-C_α vector extending away from the hydrophobic core, and hence is termed "parallel packing"; while in the **d** position, the C_α-C_β vector extends perpendicular to the C_α-C_α vector into the core, and hence is termed "perpendicular packing". As a result of these geometric differences, the contribution that a particular amino acid residue can make towards the stability of the hydrophobic core is thus dependent not only on its hydrophobicity but also on its ability to pack efficiently within its local geometrical context.

In addition to the significant contribution the interfacial **a** and **d** residues can play in determining coiled-coil stability, it has become apparent that the type and placement of these residues can also influence the associative state of the structure. For example, mutational studies in the GCN4 protein have shown that the presence of a central Asn residue in an **a** position imparts a strong tendency for the coiled-coil structure to remain exclusively in the two-stranded oligomeric state (Gonzalez *et al.*, 1996a,b,c; Lumb and Kim, 1995; O'Shea *et al.*, 1993). Monera *et. al.*, (1996) have shown that the placement of alanine residues in the central position of a model coiled-coil can regulate the difference between a two- and four-stranded structures. Also Harbury *et. al.*, (1993) using synthetic analogs of GCN4 have shown that a homogenous interface of isoleucine residues at positions **a** and **d** resulted in a trimeric coiled-coil, substitution of leucine residues at the **a** positions resulted in tetramer formation, while an interface of Ile residues at **a** and Leu residues at **d** resulted in a two-stranded dimeric state. Thus, although subtle and varied, the contribution of the **a** and **d** positions on oligomerization state can also be quite significant.

Recently, we have carried out the complete biophysical characterization of a model coiled-coil protein in which 20 different amino acids residues were substituted at an **a** position to determine their effect on protein stability . As observed by Wagschal *et al.* (1999a,b or chapters VI and VII), clear stability differences and hence amino acid residue preferences were observed between the 20 different analogs. For example, the β-branched residues of Ile and Val were indeed the most stable amino acid residues occurring in an **a** position, in excellent accord with the frequency of occurrence of such residues found within native protein coiled-coil sequences as well as their predicted preferential packing in such a position. The polar and charged residues were much less stable as expected by their hydrophobicity. Interestingly, several residues showed significant deviations from that of their expected stability based on hydrophobicity, and molecular modeling studies showed this was due in part to their ability to pack either more or less efficiently within the **a** position geometric context.

Additionally, because the model coiled-coil sequence could adopt both a two-stranded and a three-stranded association state, the oligomeric preferences of each individual amino acid were also observed. For example, Asn directed only the two-stranded oligomeric state similar to that observed for the GCN4 protein (Gonzalez *et al.*, 1996a,b,c; Lumb and Kim, 1995; O'Shea *et al.*, 1991, 1993), while Leu induced preferentially the three-stranded oligomeric state in benign buffer conditions, in accord with the results of Harbury (1993). Thus, both a relative stability scale as well as resultant oligomeric preferences of the 20 amino acid residues in the **a** position were acquired.

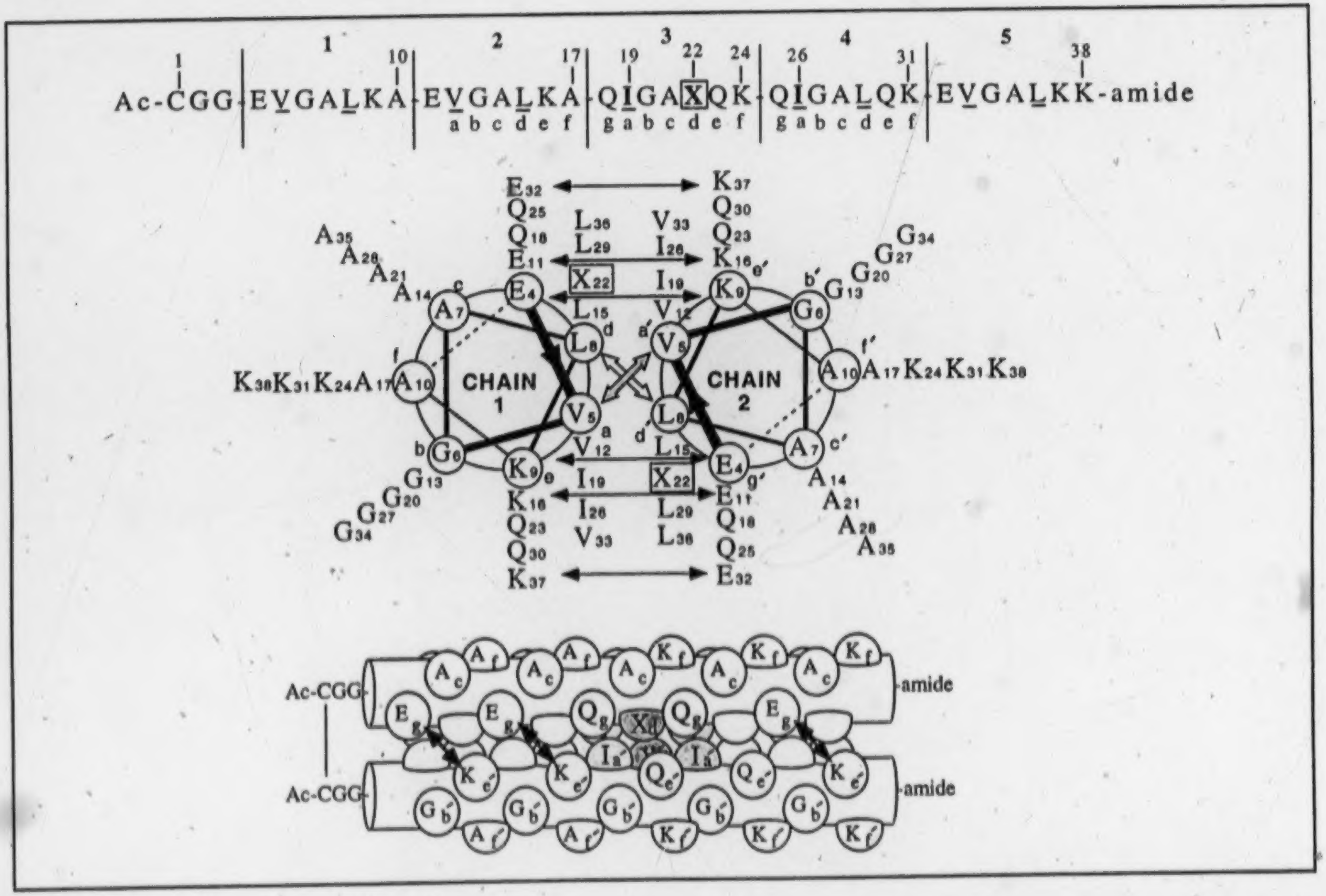
Although these results are directly amenable to further interpretation of the specific roles of various amino acid residues occurring in an **a** position of native coiled-coils, they are not directly transferable to those in the hydrophobic core **d** position, due in part to its structurally unique and different packing arrangement. Thus, the goal of the present study was to repeat the same 20 amino acid residue substitutions but this time in a hydrophobic core **d** position of a model coiled-coil protein. These results are presented below.

B. Results

*De novo design of the model **d** coiled-coil sequence*

Figure VIII-1 top, shows the amino acid sequence for one strand of the model **d** coiled-coil protein. Residues constituting the hydrophobic core **a** and **d** positions are underlined while the guest site at which the 20 different amino acid residues were substituted is boxed and denoted by an X. The model sequence consists of 38 residues, 35 of which are directly involved in forming the α-helical coiled-coil structure. A Cys-Gly-Gly linker was incorporated at the N-terminus to allow formation of an intermolecular disulfide bridge between two polypeptide chains, while the two glycine residues added conformational flexibility and distance between the coiled-coil and the disulfide bridge. The hydrophobic core of the model sequence was designed to consist largely of valine and leucine residues in the **a** and **d** positions, respectively (Ile residues at position 19 and 26 are discussed further below). The choice of these residues was based on earlier results of Wagschal *et al.* (1999a,b), which showed that these residues can generate a sufficiently stable hydrophobic core to carry out the desired substitutions, yet be structurally heterogeneous in that a hydrophobic core containing these residues can adopt either two- and three-stranded oligomeric states. Thus, the ability of guest site residue to influence the stability as well as the association state can be measured. Positions **b** and **c** of the coiled-coil contained Gly

Figure VIII-1. Model **d** coiled-coil. **Top**, Amino acid sequence of the model **d** coiled-coil peptide used in this study. Residues involved in forming the 3-4 hydrophobic repeat at position **a** and **d** are underlined. The site of amino acid substitution within the sequence is denoted by X and boxed (i.e. residue 22). Isoleucine residues, 19 and 26, which directly flank the site of substitution in the folded state are bolded. Ac- denotes N^{α}-acetyl, and amide denotes C^{α}-amide; the cysteine residue used for disulfide bond formation is located at position 1. The 5 heptads forming the 35-residue coiled-coil are numbered 1 through 5. The positions of the heptad are labelled 'gabcdef'. **Middle**; End view of the model **d** sequence starting from the NH_2 terminus at residue 4. The direction of propagation of the helices is into the page from NH_2 to COOH terminus with the polypeptide chains parallel and in-register. Residues in the first two helical turns are *circled*. Heptad positions are labeled *a-g*, with the *prime* indicating corresponding positions on the opposing helix. Wide arrows depict the hydrophobic interactions that occur between residues in the a and d positions. Solid arrows represent *i* to *i'*+5 electrostatic attractions which can occur between the "g" position of one helical strand and the "e'" position of the adjacent helical strand (e.g. positions 4 to 9', 11 to 16', and 32 to 37'). **Bottom**; Side view. Residues in positions *b* ,*c*, *e*, *f* and *g* are shown. Potential electrostatic interactions occurring across the hydrophobic face directing parallel registry are indicated by solid arrows. The site of substitution X (position 22d) within the *hydrophobic core* as well as residues which directly flank the site are *shaded*.


Ac-CGG-EVGALKA-EVGALKA-QIGAXQK-QIGALQK-EVGALKK-amide
1 2 3 4 5
1 10 17 19 22 24 26 31 38
a b c d e f g a b c d e f g a b c d e f
CHAIN 1
CHAIN 2
E32 Q25 Q18 E11 E4
K37 Q30 Q23 K16 K9
L36 L29 X22 L15 L8
V33 I26 I19 V12 V5
A35 A28 A21 A14 A7
K38 K31 K24 A17 A10
G34 G27 G20 G13 G6
a b c d e f
a' b' c' d' e' f' g'
Ac-CGG
amide
Ac Af Kf Eg Qg X Ia Ke' Qe' Gb' Af' Kf'

and Ala residues, while position **f** contained two Ala and three Lys residues. The three Lys residues were introduced to increase solubility, discourage aggregation and maintain a net positive charge on each coiled-coil strand regardless of the amino acid residue substituted at the guest site.

Position 22, a **d** position, located within the 3rd heptad was chosen as the site of substitution since perturbing the hydrophobic core in this region has been shown to have the greatest effect on coiled-coil stability and oligomerization states (Zhou *et al.*, 1992a; Harbury *et al.* 1993, Wagschal, *et al.*, 1999a,b). Moreover, the central location allows for the assessment of stability changes independent of any interactions of the various residues (particularly the charged ones) with the macro helix dipole and minimizes the helix capping effects of the various residues.

Figure VIII-1, middle and bottom, shows the helical wheel and side view diagrams of the model **d** protein, respectively. The two-headed open arrows in the hydrophobic core depict van der Waals interactions between "**a**" and "**d**" residues on adjacent chains. The two-headed solid arrows show the location of electrostatic salt bridges between Lys and Glu residues at positions "**e**" and "**g**" of adjacent α-helices (i to i'+5 or g to e') designed to stabilize the coiled-coil structure further as well as orient the polypeptide chains in an in-register and parallel manner in conjunction with the N-terminal disulfide bond.

As highlighted in Figure VIII-1, bottom, the hydrophobic interfacial environment surrounding the substitution site is defined by Ile 19 and Ile 26. Ile residues were incorporated at the adjacent '**a**' positions (as opposed to Val) in order to match isosterically the same type of hydrophobic environment as that designed in the X19**a** model coiled-coil protein (Wagschal *et al.*, 1999a) as well as mimic the most often observed structural environment in native coiled-coil sequences (Cohen and Parry, 1990; Lupas *et al.*, 1991). Additionally, Gln residues were strategically placed at positions 18**g**, 23**e**, 25**g** and 30**e** to provide a moderately polar yet uncharged micro-environment directly adjacent to the guest site, thus precluding the formation of interhelical e-g' (or e'-g) salt bridges around the substitution site and potential charge-charge electrostatic interactions between the side chain of these residues and the guest site amino acid residue side chain. Lastly, the N- and C-termini of the individual polypeptide chains were acetylated and amidated in order to eliminate charge-charge repulsions at the termini as well as possible unfavourable electrostatic interactions with the α-helix dipoles (Shoemaker *et al.*, 1987).

Secondary structure analysis

To determine the effect of each guest site amino acid residue on the secondary structure of the model **d** coiled-coil protein, each of the analogs were analyzed by far-ultraviolet circular dichroism (CD) spectroscopy (Table VIII-1). Each analog, with the exception of proline, showed a CD spectrum characteristic of α-helical proteins (Chen *et al.*, 1974) with a double minima at 208 and 222 nm and a maxima at 192 nm (data not shown). The average mean residue molar ellipticity for all analogs was -30,800 deg. cm^2 $dmol^{-1}$, indicating almost all analogs were >95% in the fully folded state as compared to the theoretical value of -34,700 degrees for a 100% α-helical 35 residue peptide (Chen *et al.* 1974). Moreover, the ratio of molar ellipticity of 222 nm/208 nm for each analog, with the exception of Arg, Orn and Pro, were greater than unity indicative of fully folded coiled-coil structures (Engle *et al.*, 1991; Lau *et al.*, 1984; Monera *et al.*, 1993; Zhou *et al.*, 1992a,b). Thus, the observation of similar molar ellipticity values and similar 222/208 ratios suggests that all of the analogs with the exception of proline can be incorporated into the model **d** coiled-coil scaffold sequence with minimal perturbation to the native structure. In the case of proline, its presence within the guest site caused essentially a complete loss in the α-helical content of the coiled-coil, indicating proline residues are completely unacceptable within a central **d** position. Even in the presence of 50% TFE, only ~50 % of the helical content could be restored. Hence, the absence of α-helical content in benign buffer conditions precluded its ability to be compared further with any of the other analogs, and is not mentioned hereafter.

Table VIII-1. Biophysical characterization of the synthetic model **d** peptides

Substituted[a] Amino Acid	$[\theta]_{222}$ [b] Benign	$[\theta]_{222}$ [b] 50%TFE	$[\theta]_{222/208}$ [c]	$[\mathrm{GdnHCl}]_{1/2}$ [d]	*m*[e]	$\Delta\Delta G_u(\mathrm{Ala})$ [f]	t_R [g]	$\Delta t_R(\mathrm{Ala})$ [h]	$t_R\#$ [i]
	degrees · cm² · dmol⁻¹			(M)		*kcal · mol⁻¹*	(min)	(min)	
Leu	-32,530	-28,820	1.04	3.76	1.54	3.8	67.4	7.8	2
Met	-33,050	-29,380	1.08	3.32	1.71	3.2	63.9	4.3	6
Ile	-32,520	-28,800	1.06	3.25	1.69	3.0	67.5	7.9	1
Tyr	-34,350	-29,020	1.04	2.31	1.98	1.4	58.8	-0.8	8
Phe	-32,300	-28,190	1.06	2.21	1.98	1.2	65.4	5.8	3
Val	-32,770	-24,720	1.06	2.20	1.74	1.1	64.5	4.9	5
Gln	-34,150	-30,700	1.06	1.90	1.94	0.5	52.3	-7.3	15
Ala	-30,610	-23,780	1.06	1.64	2.04	0.0	59.6	0.0	7
Trp	-30,350	-26,840	1.05	1.58	2.08	-0.1	64.7	5.1	4
Asn	-31,540	-20,760	1.09	1.38	2.50	-0.6	51.4	-8.2	18
His	-30,020	-25,010	1.09	1.26	2.33	-0.8	53.3	-6.3	11
Thr	-33,240	-28,240	1.06	1.17	2.94	-1.2	57.1	-2.5	9
Lys	-30,660	-26,030	1.01	0.89	2.65	-1.8	53.3	-6.3	12
Ser	-31,730	-31,260	1.06	0.82	2.64	-1.8	54.0	-5.6	10
Asp	-33,070	-23,150	1.05	0.78	2.15	-1.8	48.3	-11.3	19
Glu	-24,830	-23,400	1.01	0.32	2.10	-2.7	47.9	-11.7	20
Arg	-21,880	-23,870	0.98	0.35	2.43	-2.9	53.1	-6.5	14
Gly	-27,990	-22,890	1.06	0.31	3.34	-3.6	53.0	-6.6	13
Orn	-16,560	-25,360	0.88	0.13	2.02	-3.1	51.7	-7.9	17
Pro	-4,592	-17,510	0.49				52.4	-7.2	16

Table VIII-1 Legend

---

a Amino acid sequence for the model **d** peptide is shown in Figure VIII-1. Position 22 was substituted with each of the listed amino acids.

b The mean residue molar ellipticities at 222 nm were measured at 25°C in benign buffer (0.1 M KCl, 0.05 M PO_4, pH 7). For samples containing TFE, the above buffer was diluted 1:1 (v/v) with TFE. Peptide concentrations were 100 µM.

c The helical ratio $[\theta]_{222/208}$ was calculated by dividing the observed molar ellipticity value at 222 nm ($[\theta]_{222}$) by the observed molar ellipticity value at 208 nm ($[\theta]_{208}$).

d $[\mathrm{GdnHCl}]_{1/2}$ is the transition midpoint, the concentration of guanidine hydrochloride (M) required to give a 50% decrease in molar ellipticity at 222 nm between the folded and unfolded states.

e *m* is the slope in the equation $\Delta G_u = \Delta G_u^{H_2O} - m[\mathrm{GdnHCl}]$ with units kcal mol^{-1} M^{-1}.

f $\Delta\Delta G_u(\mathrm{Ala})$ is the difference in free energy of unfolding relative to the Ala substituted peptide. $\Delta\Delta G_u(\mathrm{Ala})$ was calculated using the equation: $\Delta\Delta G_u = ([\mathrm{denaturant}]_{1/2,X} - [\mathrm{denaturant}]_{1/2,\mathrm{Ala}})(m_X + m_{\mathrm{Ala}})/2$ (Serano and Fersht, 1989) were X denotes the other substituted amino acid analogs. Positive values denote that the substituted residue analog is more stable than that of the Ala analog, negative values denote that the substituted residue analog is less stable than that of the Ala analog. The calculated $\Delta G_u^{H_2O}$ for the Ala substituted peptide was 3.03 kcal/mol.

g Observed RP-HPLC retention time for each model **d** peptide analog at pH 6.5, see material and methods for more details.

h $\Delta t_{R(\mathrm{Ala})}$ is the difference in RP-HPLC retention time of each peptide analog relative to the Ala substituted peptide. Positive values indicate an increase in retention time relative to the Ala analog.

i Ranked order of hydrophobicity of the 20 amino acid analogs based on RP-HPLC, 1 denotes most hydrophobic, 20 least hydrophobic.

Other notable differences in the helicity values observed include the charged residues (Table VIII-1). In the case of the positively charged side-chain residues of Lys, Arg and Orn, the Lys analog displayed essentially a fully helical structure. Movement of the positive charge closer to the hydrophobic core by one methylene group, as with Arg and Orn, resulted in significant decreases in the helicity of these analogs (69% and 52% of the Lys analog, respectively), suggesting either unfavourable interactions with the core or a higher energetic cost of their desolvation near the core. Interestingly however, this trend was not strictly adhered to in the case of the negatively charged residues. The Glu substituted analog was ~ 75% of the predicted theoretical helical content, while movement of the charged carboxyl group one methylene group closer to the core, as in the case of the Asp analog, resulted in a fully helical state. Thus the fully folded state of the Asp analog vs. Glu suggests favorable structural interactions are present and are offsetting the closeness of charged carboxyl group to the hydrophobic core.

In the presence of TFE, all of the model **d** analogs, with the exception of Arg or Orn, showed a decrease in the 222 nm reading and a switch in the 222/208 nm ratio to ~0.8 indicating that all of the structures are making a transition from a two- or three-stranded folded state in benign buffer to a single stranded α-helical state (Cooper and Woody, 1990; Lau *et al.*, 1984, 1990; Sönnichsen *et al.*, 1992). The significant variation in the TFE values, e.g., those of Asn and Asp which were ~10,000° less than in benign medium, indicates that these residues are clearly more dependent on the packing of the helices to remain in a helical conformation than when in a single stranded state.

*Guanidine-HCl denaturation studies of the model **d** analogs*

To determine the effect each guest site residue on the stability of the coiled-coil structure, each model **d** coiled-coil protein was denatured using GdnHCl and the ellipticity at 222 nm, reflective of the α-helical content of the structure, monitored by CD spectroscopy. Figure VIII-2, panels A and B, show overlay plots of the GdnHCl

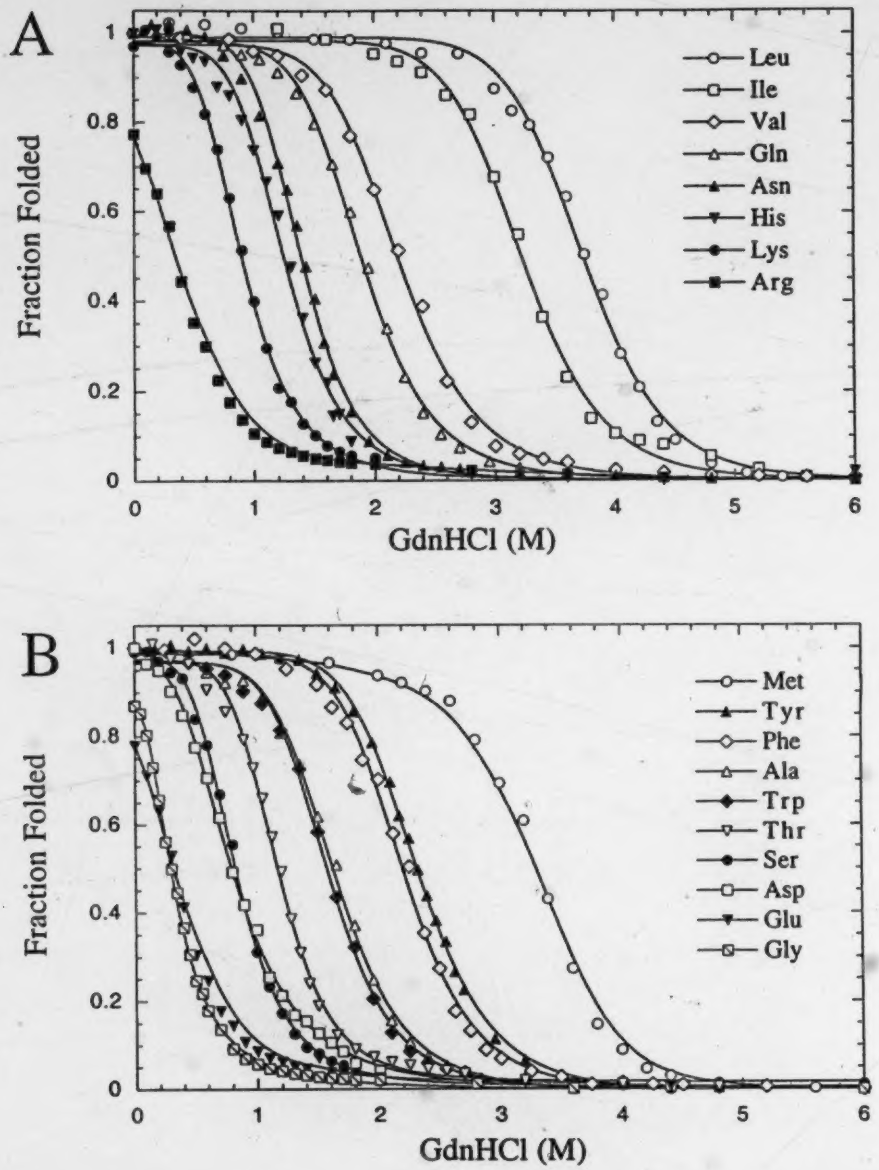


Figure VIII-2. Panels A and B; GdnHCl denaturation profiles of the model d peptides. Denaturations were carried out at 25 °C in 0.1 M KCl, 0.05 M PO_4, pH 7 buffer with GdnHCl denaturant. The fraction folded (*ff*) of each peptide was calculated as described in the Materials and Methods section. Each peptide was analyzed at 100 μM concentration.

denaturation profiles for 18 of the model **d** analogs. As observed, all of the analogs displayed a denaturation profile which could be easily fit by a sigmoidal equation indicative of a two-state single transition cooperative unfolding process. Comparison of the GdnHCl midpoints (see Table VIII-1) for each transition show that although the mean residue molar ellipticity values were similar under benign buffer conditions, the resultant substitution caused a great variation in the overall stability of the model **d** coiled-coil. In general, the hydrophobic amino acid analogs Leu, Met, Ile, Tyr, Phe and Val showed the greatest stability, these were then followed by the polar residue analogs of Gln, Asn, Thr and Ser, and finally by the charged residue analogs of Lys, His, Asp, Glu and Arg. Two analogs were notably different than their expected stability, these were Trp, which clearly demonstrated a lower than expected stability for its size and hydrophobic character (i.e. it was similar to the small hydrophobic sidechain Ala), and Asp, which was significantly more stable than the Glu analog despite the Glu analog having a slightly greater hydrophobicity due to an extra methylene group in its sidechain.

Hydrophobicity

Stability of coiled-coils has classically been related to the hydrophobicity of the side chains existing in the hydrophobic core **a** and **d** positions. To compare and contrast the differences between the observed stability and its hydrophobicity in the corresponding **d** position, we determined the hydrophobicity of each of the model **d** peptide analogs using reversed-phase high-performance liquid chromatography (RP-HPLC) at pH 7. A summary of the observed elution times (t_R), their 'relative elution times' compared to that of the Ala analog (Δt_R(Ala)), and their relative ranked order from most hydrophobic to least hydrophobic is shown in Table VIII-1, columns 8, 9, and 10, respectively. The RP-HPLC method conditions employed (see methods for more details) resulted in all 20 analogs being well resolved with a 19.5 minute difference in retention time between the most hydrophilic analog (Glu t_R = 47.9 min) and the most hydrophobic analog (Leu t_R = 67.4 min). The

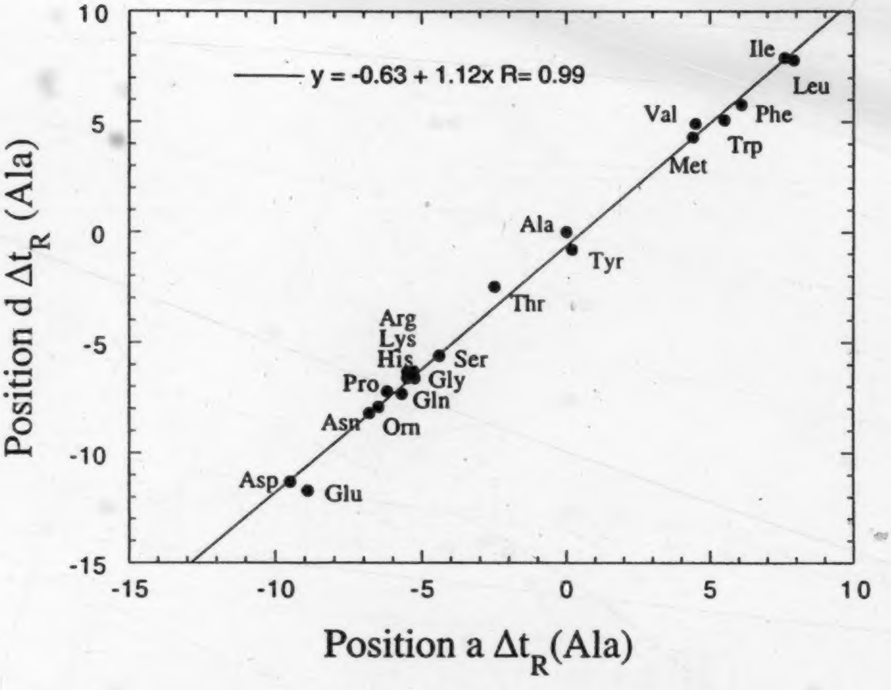


Figure VIII-3. A plot of the relative RP-HPLC retention times of the monomeric model **d** peptide analogs versus those obtained from the model **a** peptide sequence (Wagschall *et al.* 1999).The cysteine side-chain (position 1) was alkylated with iodoacetamide to prevent disulfide bridge formation. The change in retention time (Δt_R) relative to the alanine analog is plotted for the "a" and "d" positions. Least squares fit analysis shows R = 0.998. For RP-HPLC conditions see materials and methods section.

relative hydrophobicity values of the 20 amino acids, as expressed by the RP-HPLC retention times of the 20 analogs, appeared to be essentially the same as those reported previously using octanol water partitioning (Fauchere & Pliska, 1983; Eisenberg and McLachlan, 1986) as well as other model systems including our own X19a model peptide sequence (Chapters VI and VII; Wagschal *et al.* 1999b and references therein). For example, a plot between the observed Δt_R(Ala) times of this study and those obtained in the previous study (Δt_R(Ala) **a** position) for the X19**a** sequence showed a correlation of $r = 0.99$ (Fig. VIII-3A) thus indicating that the hydrophobicity values derived for each side chain are largely an intrinsic parameter of the side-chain and not model or position dependent .

Oligomerization

To assess whether the guest site residue influences the association state of the model **d** coiled-coil protein in a manner similar to that previously observed for the X19**a** coiled-coil protein (Wagschal *et al.*, 1999a,b), each of the 20 model **d** analogs were analyzed by high performance size-exclusion chromatography (HPSEC). Previously, we demonstrated that this technique (HPSEC) offers a reliable and quick approach for establishing the oligomerization state as well as respective ratios between monomer and higher order oligomeric states, if present. This is due in part to the observation that such associative states exchange relatively slowly in time and thus allow for their respective separations during the chromatographic run. HPSEC runs were carried out by injecting each analog in benign buffer from either a stock solution of 500 μM or a 10 fold diluted sample of 50 μM (i.e., similar in concentration to that used for CD measurements). Figure VIII-4, panel A, shows representative HPSEC chromatograms for the model **d** analogs Val, Glu and Ser. In general, the model **d** analogs eluted from the size-exclusion column as either single peaks with molecular weights corresponding to the two-stranded disulfide bridged monomer (~7,000 Da), the three stranded trimer (~21,000 Da) or mixtures of the two (e.g. Glu, Val

---

Figure VIII-4. *Panel A*, Representative HPLC-size exclusion chromatograms for the model **d** peptide analogs Val, Glu and Ser. Each analog was run in benign buffer (0.1 M KCl, 0.05 M PO_4, pH 7) at a flow rate of 0.2 ml/min on a Sephadex 75 column. M and T denote monomeric and trimeric molecular weight species. Each run consisted of a 100 μl injection of a 0.5 mg/ml solution of the corresponding peptide. *Panel B*, Representative HPLC-size exclusion chromatograms for the model D peptide analog Ser in the presence of increasing GdnHCl concentrations . Each analysis was run in benign buffer 0.1 M KCl, 0.05 M PO_4, pH 7 containing 0.0 to 2.0 M GdnHCl as denoted in the figure. T_f denotes folded trimer, M_f and M_u denote the two-stranded disulfide bridge folded and unfolded monomer, receptively. Each run consisted of a 100 μl injection of a 0.5 mg/ml solution of the Ser peptide analog dissolved in the running buffer. All runs were performed at ambient temperature.

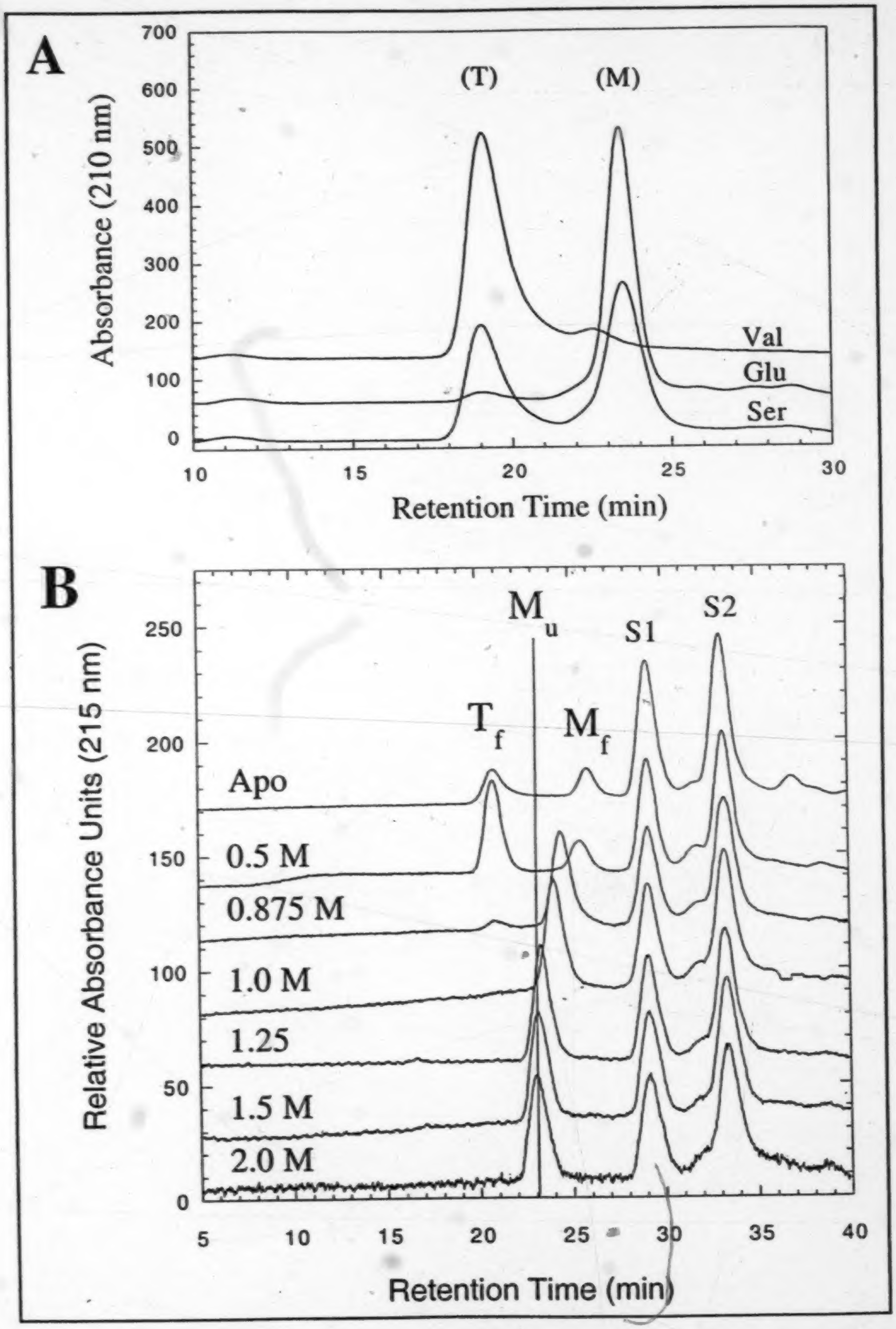

A
(T)
(M)
Absorbance (210 nm)
700
600
500
400
300
200
100
0
Val
Glu
Ser
10
15
20
25
30
Retention Time (min)
B
M_u
S1
S2
T_f
M_f
Apo
0.5 M
0.875 M
1.0 M
1.25
1.5 M
2.0 M
Relative Absorbance Units (215 nm)
250
200
150
100
50
0
5
10
15
20
25
30
35
40
Retention Time (min)

Table VIII-2 Oligomerization Data.

Substituted[a] Amino Acid	Position a[b]		Position d[c]	
	% Trimer[d]	% Monomer[d]	% Trimer[d]	% Monomer[d]
Leu	**100**	0	31	69
Ile	39	61	**100**	0
Val	57	43	**100**	0
Met	90	10	80	20
Ala	38	62	39	61
Gly	35	65	50	50
Phe	65	35	46	54
Tyr	**100**	0	0	**100**
Trp	0	**100**	22	78
Thr	76	24	**90**	10
Ser	25	75	42	58
Gln	**100**	0	39	61
Asn	0	**100**	72	28
Lys	0	**100**	0	**100**
Orn	0	**100**	0	**100**
Arg	0	**100**	0	**100**
His	**100**	0	45	55
Glu	46	54	0	**100**
Asp	0	**100***	0	**100**
Pro	----	----	----	----

Table VIII-2 Legend

---

a Amino acid residue substituted at position 22d of the model **d** peptide sequence (Figure VIII-1) or at position 19a of the model **a** peptide sequence (Wagshal et al 1999a,b).

b Observed percent oligomerization states of the model **a** peptide analogs at 50μM peptide concentration in benign buffer (0.1 M KCl, 0.05 M PO4, pH 7) using HPSEC (Wagschal et al. 1999a,b). Residues which displayed only a single oligomeric state are bolded.

c Observed percent oligomerization states of the model **d** peptide analogs at 50μM peptide concentration as observed by HPSEC chromatography in this study (see Materials and Methods for more details). Residues which display only one type of oligomeric state are bolded.

d Monomer is the disulfide bridged two-stranded α-helical coiled-coil (~7000 daltons) and the trimer is a three stranded coiled-coil consisting of three disulfide bridged peptides (~21,000 daltons) as depicted in Wagschal et al 1999a,b.

* The Asp analog was 100% monomeric at 5ºC, and 50% monomeric, 50% unfolded at 25ºC.

and Ser, respectively, Fig. VIII-4A). No peaks greater or intermediate in apparent molecular weight were observed (i.e. dimers or tetramers). The observation of monomeric and/or trimeric states depending on the amino acid substituted indicates, firstly, that the model **d** coiled-coil sequence is indeed similar to that of the model X19**a** coiled-coil sequence in that a hydrophobic core of three Val and two Ile in the "**a**" position and four Leu in the "**d**" position is structurally heterogeneous such that this core can adopt both a two-stranded monomeric or a three-stranded trimeric oligomeric state; and secondly, that a single residue substituted at a **d** position can influence the final association state. Thus, the effects of each of the 20 amino acid side chains in their ability (or inability) to define a single oligomeric state can be assessed. Table VIII-2 shows a summary of the integrated areas of each eluted peak as a percentage of total area generated by each protein observed by HPSEC, adjacent to those previously reported for the same substitutions in the X19**a** model coiled-coil protein (Wagschal *et al.*, 1999a,b). As observed in Table VIII-2, all of the charged residues with the exception of His clearly defined only the monomeric two-stranded state. Even at a 10 fold higher injected concentration (data not shown), no higher order oligomeric states were observed for these analogs, suggesting that the presence of the charged group clearly destabilizes the three-stranded state. In contrast, the β-branched residues of Ile, Val and Thr demonstrated a strong tendency to populate exclusively the three-stranded trimeric oligomeric state, suggesting that the greater burial of surface area in an "acute" packing arrangement (vs. perpendicular) is apparently more favorable for these residues in the three-stranded state vs. that of the two-stranded monomeric state. Interestingly, several analogs, His, Asn, Gln, Ser, Trp, Tyr, Phe, Gly, Ala and Leu, were indiscriminate in their ability to define a single oligomeric state and thus adopted various ratios of the monomeric and trimeric states depending on the concentration. It is noteworthy that of these analogs the Asn analog which clearly denoted only the two-stranded monomeric state in the "**a**" position is completely devoid of this effect in the "**d**" position .

To reinforce the accuracy of the HPSEC data, several of the model **d** analogs were analyzed by sedimentation equilibrium (SE) ultracentrifugation. Since this technique does not suffer from shape dependence or possible matrix interactions as does size-exclusion chromatography, a more accurate determination of molecular weight in solution is gained. Representative curve fits corresponding to Val, Glu and Lys are shown in Figure VIII-5 . As observed by HPSEC the Val data were best fit to a single species fit for a three-stranded associated model, the Glu data were best fit to a single species fit for a two-stranded monomeric model, while the Lys data were best fit to a monomer to trimer equilibrium fit. Thus the definition of oligomeric states and their relative variation for these peptides by sedimentation equilibrium analysis appear to be in excellent agreement with the results observed by HPSEC. Moreover, SE data generated for several other analogs (data not shown) were also in good agreement with that observed by HPSEC, thus strongly reinforcing the HPSEC data.

Trimer to monomer switch.

A critical issue in being able to compare directly the stability data for each of the model **d** analogs is that each of the analogs is required to be denatured from the same oligomeric state. Previously, we showed that in the case of the model X19**a** analogs (Wagschal *et al.*, 1999a,b), that the variation in associative states observed initially in benign buffer medium were eliminated upon addition of denaturant. That is, all peptides in the trimer state or mixtures of trimers/monomer states changed to fully folded two-stranded monomeric states prior to unfolding. To establish if this is also the case for the model **d** analogs, several of the heterogeneous and trimeric analogs where subjected to HPSEC in the presence of increasing concentrations of GdnHCl prior to and through their denaturation transition areas. A representative series of elution profiles for the Ser analog is shown in Figure VIII-4B. Similar to that observed for the model X19**a** analogs, the Ser analog (initially occurring in ~50% trimer:monomer) completely reverts to the two-stranded monomeric state at a

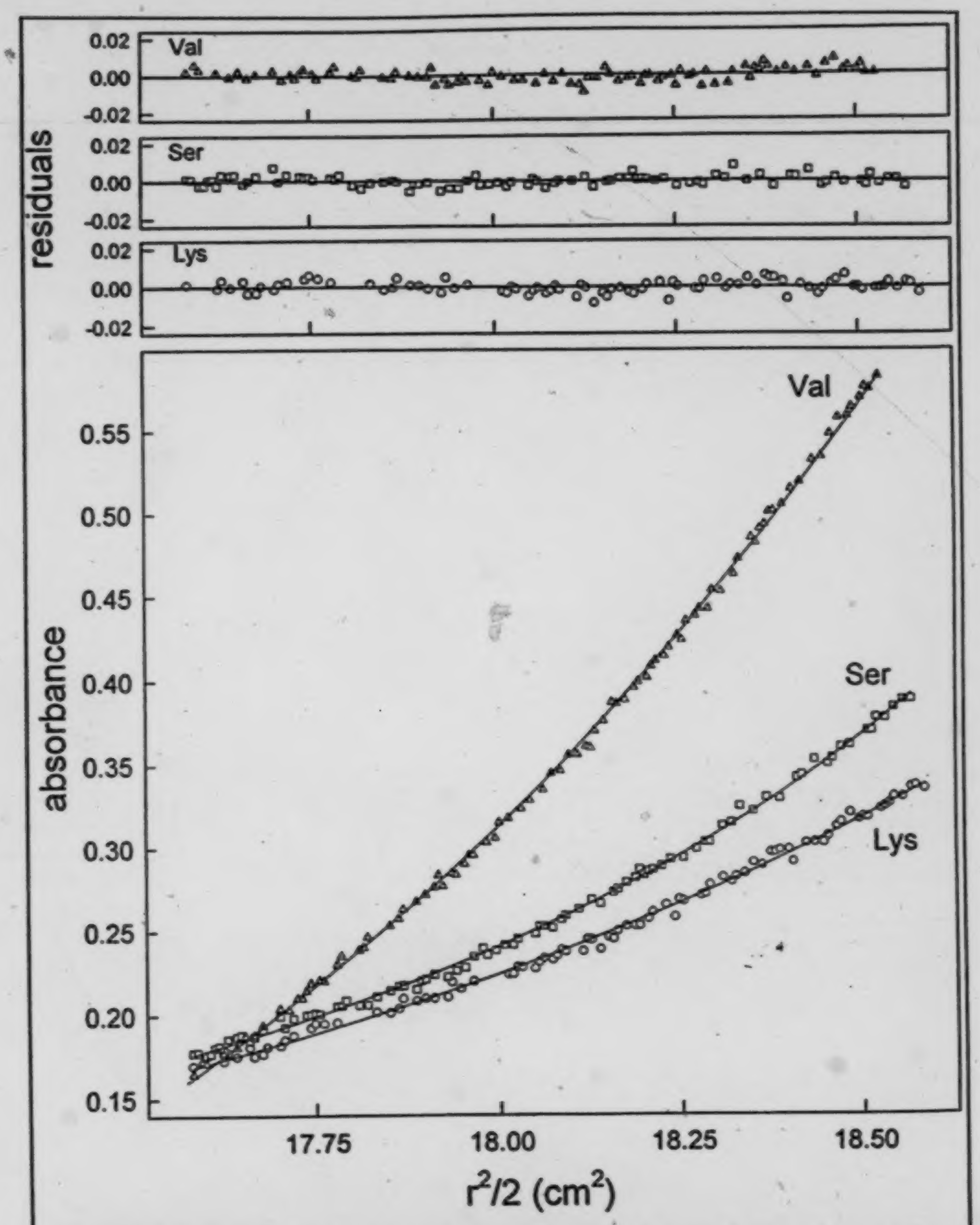


Figure VIII-5. A plot of concentration versus radial distance from sedimentation equilibrium analysis of the model **d** analogs Val, Ser and Lys peptides performed at 18,000 and 26,000 rpm. The Val data was best fit to a single species model with a trimeric molecular weight (~22,000 Da). The Ser data was best fit to an associating monomer to trimer equilibrium model (~14,500 Da). The Lys data was best fit to a single species model with a monomeric molecular weight (~7,500 Da). Plots of the residuals from each curve fit are shown above, respectively.

GdnHCl concentration between 0.5 M and 0.875 M, at the beginning of its unfolding transition area. As the GdnHCl concentration is further increased up to 2.0 M, the retention time of the Ser analog peak gradually decreases until the retention time of the unfolded monomer state (denoted M_u) is reached. A plot of the change in retention time of the monomeric peak versus GdnHCl concentration (data not shown) shows a transition midpoint value of ~0.8 M, in excellent agreement with that observed by CD spectroscopy. Similar results to these were also observed for several of the other model **d** analogs which populated various amounts of the trimeric species. Thus, all of the heterogeneous or trimeric analogs denature from the monomeric state.

It is noteworthy that although the primary basis of the above investigation was to verify that the higher oligomeric species reverted back to monomeric species, several of the monomeric analogs were also tested to verify that they did not revert to a higher order oligomeric species upon denaturation, and no such change was observed.

*Calculation of the relative stability difference ($\Delta\Delta G_u$ of Unfolding) between model **d** analogs and their comparison to hydrophobicity.*

Having established that each model **d** analog adopts a fully folded monomeric state prior to its denaturant transition area irrespective of its initial oligomerization state in benign buffer conditions thus allowed us to treat all of the GdnHCl denaturation data similarly and calculate the difference in free energy of unfolding for each of the analogs using an equation which is concentration independent (see Methods for more detail).

Table VIII-1 shows the calculated relative stability order in terms of the difference in free energy of unfolding of each of the model **d** analogs compared to that of the substituted alanine analog (i.e. $\Delta\Delta G_{u(Ala)}$). As shown, Leu is the most stabile analog with a $\Delta\Delta G_{u(Ala)}$ of 3.8 kcal/mol (or 1.9 kcal/mol/Leu). This is followed by Met and Ile, the aromatic residues Tyr and Phe, and finally by Val and Gln. Analogs which are of lower stability than Ala, include in order of decreasing stability, Trp, Asn, His, Thr, Lys, Ser, Asp, Glu, Arg, Orn and Gly. The variation in stability includes a free energy change of 7.4 kcal/mol from the most stable to the least stable analog and thus clearly demonstrates the importance of the '**d**' position in modulating the overall protein stability. Further, one can see, packing differences in addition to hydrophobicity are clearly playing a role in defining the order. For example, the Met analog which is significantly lower in hydrophobicity than Ile (Table VIII-1 and Fig. VIII-3) is the second most stable residue in the **d** position; Tyr is slightly more stable than Phe even though Phe is significantly more hydrophobic than Tyr; Trp, which is significantly more hydrophobic than Ala, is less stable than Ala; and Asp which is significantly more stable than Glu. Moreover, only three residues (Leu, Met and Ile) are able to generate stability greater than 2.5 kcal mol^{-1} versus 5 residues previously observed in the **a** position . In general, the observed order does, however, appear to demonstrate the trend that the hydrophobic residues are more stable than the polar residues, which in turn are more stable than the charged residues, and thus fits with the general proposal that hydrophobicity is the major factor in stabilizing proteins. It is interesting to note that despite the large energetic cost associated with desolvation when bringing a charged side-chain close to the hydrophobic core interface, the absence of a side chain, i.e., creation of a cavity, caused the largest destabilization to the coiled-coil structure i.e., Gly is the least stable analog represented.

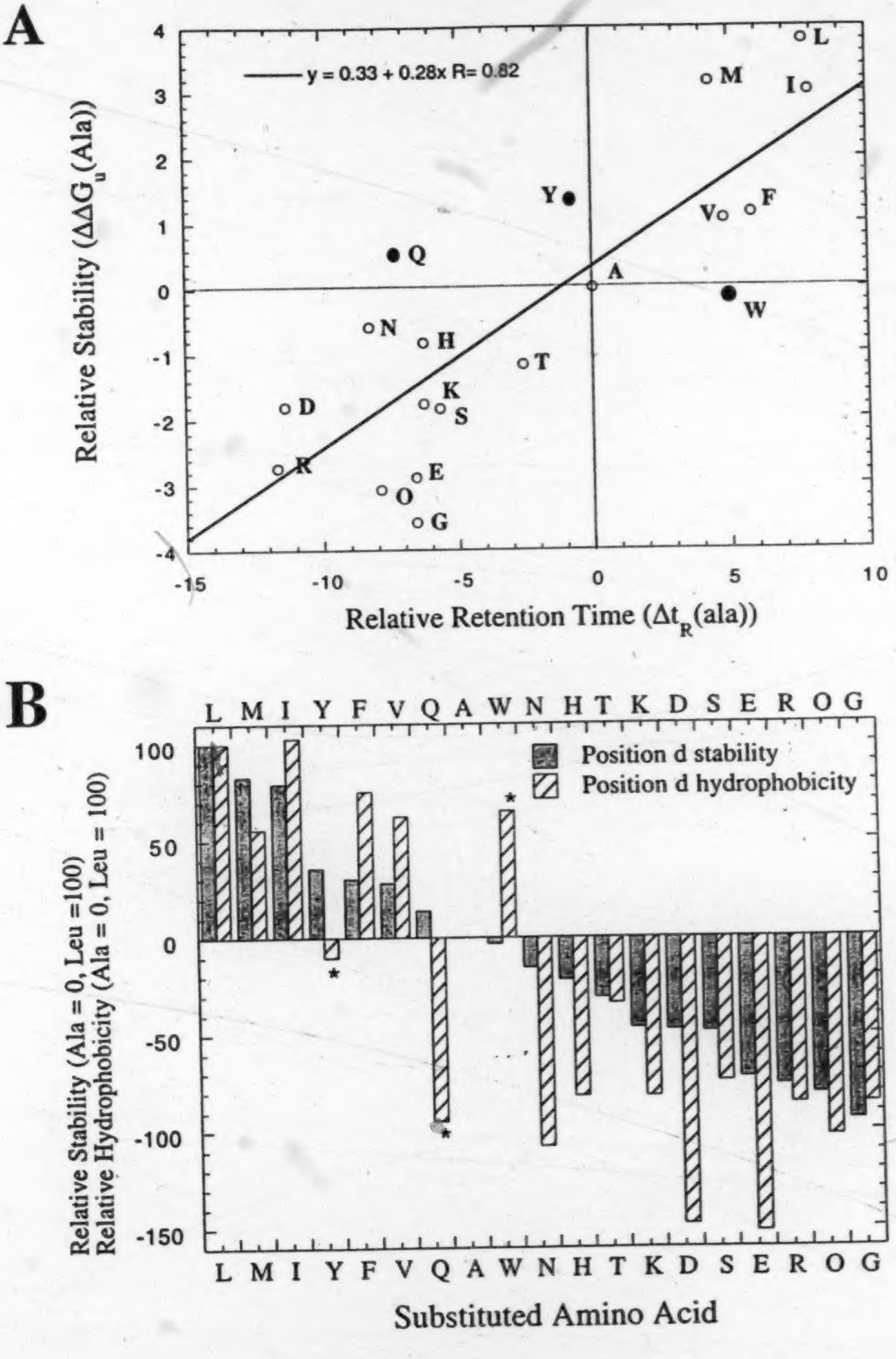
In addition to the wide variation in stability values Table VII-1 also shows that each of the model **d** position analog also exhibited significantly different slope or '*m*' value in the plot of ΔG versus denaturant concentration. In general, the least stable analogs exhibited the greatest *m* values (~2.5-3.3 kcal mol^{-1} M^{-1}) while the more stable mutants exhibited lowest values (~1.5-2.0 kcal mol^{-1} M^{-1}). The significance of these variations in terms of the folded and unfolded states are discussed in further detail below.

Stability versus Hydrophobicity

Figure VIII-6, Panel A shows a comparison between hydrophobicity of the model **d** analog with its corresponding stability. As indicated in the plot, there is in general trend, but not a good direct correlation between the two parameters (correlation coefficient 0.82). In particular, three analogs show significant deviations from their expected hydrophobicity quadrant (see legend for more details), these are the Gln, Tyr and Trp . The Trp analog, which is significantly more hydrophobic than the Ala analog, is less stable than Ala or the rest of the hydrophobic group. Tyr, which has lower hydrophobicity than Ala, is approximately equivalent in stability to Phe and Val of the hydrophobic group, and Gln, a member of the polar group, is clearly more stable than any other member of its group. Such deviations are likely representative of at least three types of packing factors, i.e., "poor packing" in the case of Trp where the complete hydrophobic side-chain is likely not buried within the hydrophobic core which lowers its stability relative to its inherent hydrophobicity; "good packing", where the polar hydroxyl group of Tyr remains solvent accessible while its aromatic ring is buried in the hydrophobic core, thus making it similar to Phe despite its relative hydrophobicity being lower, and finally, "good packing due to hydrogen bonding" which is likely in the case of Gln and Asp analogs which increases their stability relative to their hydrophobicity.

To highlight better the differences between relative hydrophobicity and stability of all of the analogs, the respective data was plotted in a histogram format where the Leu analog represents a value score of 100 and Ala represents a value score of 0 (Fig. VIII-6B). As shown, Met, the second most stable amino acid residue in the **d** position, is clearly more stable than its hydrophobicity would have indicated, suggesting favourable packing effects are clearly playing a role in defining its position in the relative stability order. The Ile and Val substituted analogs are less stable than expected based the hydrophobicity of Ile and Val, likely due to the disfavored preference of β-branched residues in the **d** position due to their less than optimal van der Waals packing. Tyr (as mentioned above) is more stable than

Figure VIII-6. Comparison of the relative stability of each substituted amino acid with its relative hydrophobicity for the "**d**" position analogs. Panel **A**; correlation plot for relative stability versus relative hydrophobicity. The upper right hand quadrant (*shaded*) represents residues which are of greater hydrophobicity and stability than that of Ala (i.e. the "nonpolar residues"). The lower left quadrant (*shaded*) represent residues which are of lesser hydrophobicity than that of Ala (i.e. the "polar/charged residues"). Residues which deviate from either group due to differences in either their hydrophobicity or stability are shown with filled points. Panel **B**; Histogram displaying the relative stability and hydrophobicity of each peptide analog. Amino acid residues indicated along the X axis are listed in order of their observed stability from greatest to least. Stability and hydrophobicity values are normalized where Leu = 100. The three model **d** peptide analogs: Gln, Tyr and Trp which appeared to deviate the most in the plot shown in panel B are indicated with an asterisk (*).



A
y = 0.33 + 0.28x R= 0.82
Relative Stability (ΔΔG_u (Ala))
Relative Retention Time (Δt_R(ala))
L
M
I
Y
V
F
Q
A
W
N
H
T
K
S
D
R
E
O
G
4
3
2
1
0
-1
-2
-3
-4
-15
-10
-5
0
5
10
B
L M I Y F V Q A W N H T K D S E R O G
Position d stability
Position d hydrophobicity
Relative Stability (Ala = 0, Leu =100)
Relative Hydrophobicity (Ala = 0, Leu = 100)
100
50
0
-50
-100
-150
*
*
*
Substituted Amino Acid

its hydrophobicity, but interestingly is approximately equivalent to that of Phe, suggesting that packing of the aromatic rings may be similar or at least independent of the para-hydroxyl group. The Trp analog is significantly less stable than expected from Trp's hydrophobicity, while the Gln and Asn analogs are clearly more stable than expected from the hydrophobicity of these residues. It is interesting to note that Gln is the more stable analog in the **d** position as compared to the **a** position where Asn was more stable (Wagschal *et al.*, 1999a,b). Finally, the charged residues (His, Lys, Asp and Glu), with the exception of Arg, are all significantly more stable than their hydrophobicity would have suggested. Particularly notable within this group is the observation that Asp is more stable than that of Glu despite Glu containing one more methylene group (as mentioned above).

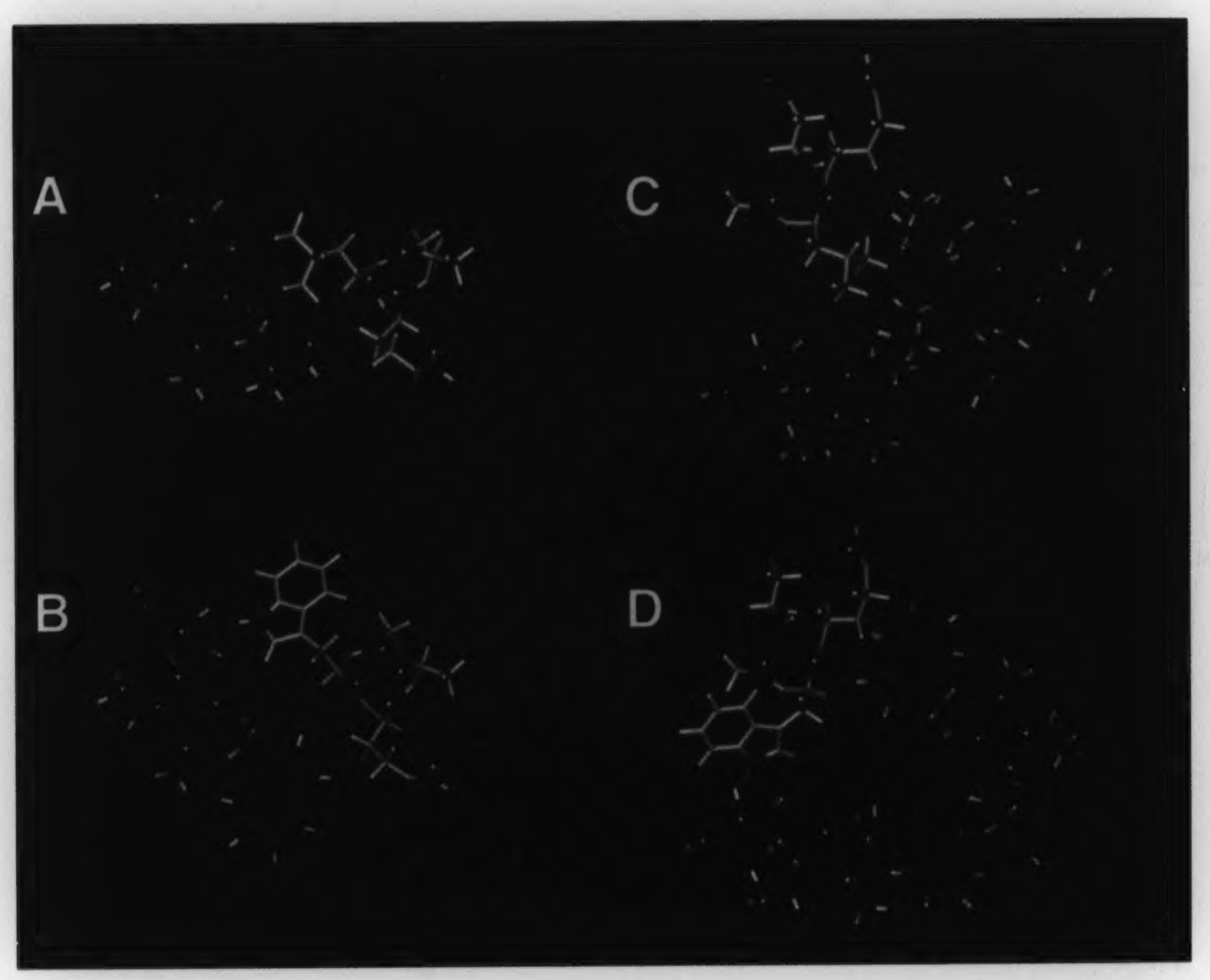
Structural analysis of Met, Trp, and Asp.

To investigate further the structural features which give rise to the apparent deviations between hydrophobicity and stability of the model **d** analogs, molecular modeling studies of each of the 20 analogs were carried out. A complete description of these results are to be presented elsewhere (Lavigne *et al., 1999*, manuscript in preparation), but for the purpose of this chapter, structural analysis of three residues (Met, Trp, and Asp) which clearly demonstrate "good", "poor" and "hydrogen bonding" packing effects are reported.

In the case of methionine, 3 crystal structures are presently available in which methionine residues can be observed in a **d** position of a homodimeric parallel two-stranded coiled-coil (pdb# 3GAP, 1HLO and 1CGP). As observed within the crystal structures as well as in the molecular modeling studies carried out herein (Fig. VIII-7A), the methionine side chain can easily adopt dihedral angles which, although not identical to leucine, can easily emulate it in creating excellent geometric complementarity and van der Waals packing interactions with the adjacent **d' a**, **a'**, **e'** and **g** residues. Additionally, methionine side chains can adopt two different preferred packing arrangements in the core position, suggesting the side-chain can possibly flip between these two arrangements and thus maintain a significant degree of

---

Figure VIII-7. Molecular models depicting the sidechain packing orientation of Met, Trp and Asp residues in a **d** position of a two stranded parallel coiled-coil. Top, Methionine sidechains are represented in stick format with the carbon atoms colored green, oxygen atoms in red, nitrogen atoms in blue, and sulfur atoms in yellow. Backbone helices are depicted in ribbon format (colored white; right side). van der Waals surface area is represented as dotted spheres (colored yellow and white for each helix, respectively, left side). Middle; the Trp sidechain. Ribbon and stick format is shown on the right, and van der Waals interactions are emphasized on the left. Bottom; molecular modeling of the Asp sidechain. The Asp sidechain is represented in stick format as above. Also shown is the predicted packing orientation of the Gln residue sidechains which exist in the adjacent helix 'e' positions. Potential hydrogen bonds between the oxygen atoms of the Asp carboxyl group and the hydrogen atom of the Gln amide sidechain group are denoted by dashed white lines.


A
C
B
D

conformational entropy in addition to its van der Waals interactionsading to stability. Further, the side-chain packing arrangement also brings the sulfur atoms within the side-chain close to the core. Because sulfur has greater polarizability than that of carbon, this would likely create greater van der Waals interactive strengths between the side chains and the core (for a more detailed explanation of the unique roles of methionine, see Nelson and Chazin (1998), and references therein). Thus, the higher than expected stability of the Met analog appears to be due to its unique ability of Met to adopt excellent geometric packing, conformational freedom and van der Waals interactions.

Molecular modeling of Trp, on the other hand, shows quite the opposite. As opposed to being buried within the hydrophobic core, the side chain of Trp extends directly out of the core (Fig. VIII-7B). Although this side chain conformation does not necessarily create a cavity in the core, it is clearly unable to pack its largest hydrophobic part, the indole ring portion, of the side-chain into the core, thus leaving it exposed to the solvent, and hence unable to be utilized. The modeling of Trp is further reinforced by the high resolution X-ray crystallographic structure of dimeric rat brain kinesin which shows two Trp residues in **d** positions of the homo two-stranded coiled-coil neck region (pdb# 3kin). In each case, the Trp residues can adopt either a parallel or perpendicular orientation relative to the helical axis of the coiled-coil, but in both of these arrangements, the large indole ring is still largely solvent exposed. Thus, the lower than expected stability of the Trp analog relative to its hydrophobicity is clearly a result of its poor "steric" packing in a hydrophobic core **d** position.

At present, no high resolution X-ray crystal structures are available for an Asp residue in either an **a** or **d** position of two-stranded coiled-coils, although it is known that such amino acid residues do occur in such positions. For example, there is one Asp residue in a **d** position of tropomyosin. Molecular modeling studies of the Asp analog show that the side chain can be packed entirely within the hydrophobic core environment with the buried carboxyl groups making intermolecular hydrogen bonds with the adjacent **g'** position Gln residue (Fig. VIII-7C). Thus, the high cost of desolvation of the carboxyl groups appears to be compensated (or offset) in part due to the formation of new intermolecular hydrogen bonds with the adjacent positions. Interestingly, the packing arrangement of the Asp side chain is almost completely isosteric with that of the packing arrangement of Leu, with the exception that the δ methyl group is replaced by a carbonyl group. Thus Asp's unique stability (i.e., greater than Glu) in a **d** position can be ascribed to its unique ability to maximize the use of its methylene component in the hydrophobic core and its hydrophilic component through hydrogen bonding adjacent to the core.

*Protein Stability versus natural occurrence at position "**d**".*

Having derived a relative stability scale for the 20 amino acid analogs, we next compared this order to the apparent frequency of occurrence for each amino acid residue occurring in a **d** position of naturally occurring coiled-coils (Cohen and Parry, 1990; Lupas *et al.*, 1991). As observed in Figure VIII-8 and Table VIII-3, Leu, which represents the most stable analog, also occurs with the greatest statistical frequency in the **d** position of native coiled-coils. Similarly, Arg and Gly, which were the least stable analogs, occurred least often. Comparison of the 17 other residues, however, indicates that in general there is no direct relationship between stability and frequency of occurrence. For example, Ala, which was only intermediate in stability, occurred with the second greatest frequency. Owing that Ala does not specifically discriminate a particular oligomeric state, *vida infra*, its high frequency in comparison with the other residues must be due inpart to its ability to modulate the coiled-coil stability without affecting the choice of two- or three-stranded coiled-coils. The significant deviations of the other residues occurrence (lower occurrence than their stability) may be due in part to the high metabolic cost associated with their biosynthesis (e.g., Trp) or their ability, or inability, to control the association state . In general, the lack of a clear correlation between stability and frequency of occurrence for many of the analogs clearly underscores the importance of the stability data obtained above

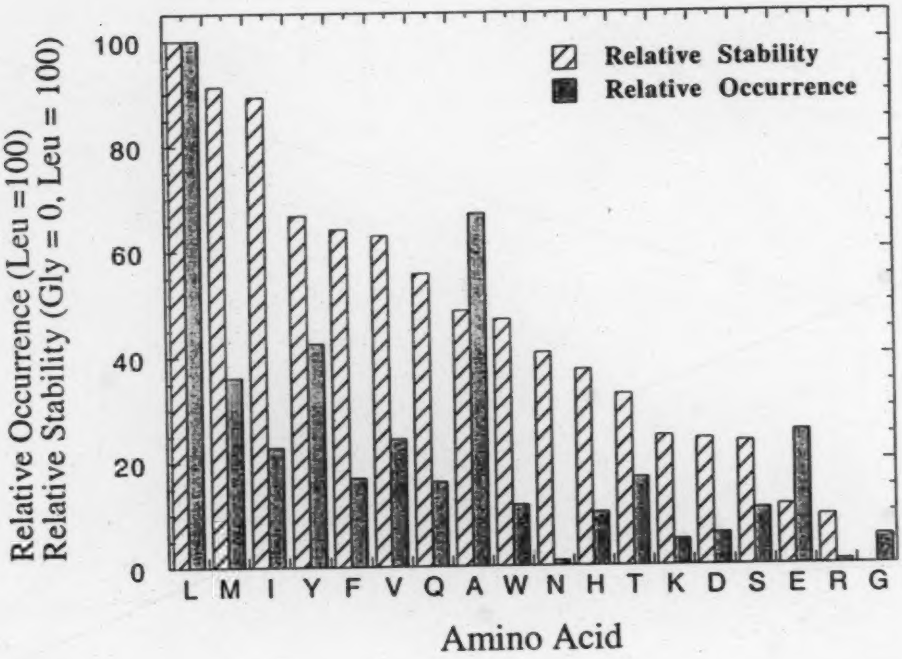


Figure VIII-8. Histograms comparing the measured relative stability and relative statistical occurrence (Lupas *et al.*, 1991) of each amino acid residue in a "d" position. Stability and statistical occurrence data have been normalized such that Leu equals 100 and Gly equals 0 for the stability data (see Table VIII-2 and legend for more details).

Table VIII-3. Summary of Stability and Statistical Occurrence Data.

Amino[a]	Position a			Position d		
Acid	$\Delta\Delta G_u$(Ala)[b]	Normalized[c] Stability	Normalized[d] Occurrence	$\Delta\Delta G_u$(Ala)[b]	Normalized[c] Stability	Normalized[d] Occurrence
Leu	3.5	100	100	3.8	100	100
Met	3.4	98	70	3.2	91	36
Ile	3.9	105	82	3.0	89	23
Tyr	2.18	74	45	1.4	67	43
Phe	3.0	89	17	1.2	64	17
Val	4.1	108	53	1.1	63	24
Gln	-0.1	41	6	0.5	56	16
Ala	0.0	43	41	0.0	49	67
Trp	0.8	55	8	-0.1	47	12
Asn	0.9	56	26	-0.6	41	1
His	-1.2	28	11	-0.8	37	10
Thr	0.2	50	5	-1.2	33	17
Lys	-0.4	37	43	-1.8	25	5
Asp	---	---	1	-1.8	24	6
Ser	-1.3	23	12	-1.8	24	11
Glu	-2.0	10	8	-2.7	12	26
Arg	-0.8	31	21	-2.9	9	1
Orn	-1.9	10	---	-3.1	7	---
Gly	-2.5	0	1	-3.6	0	6
Pro	---	---	0	----	---	0

Table VIII-3 Legend

---

a Amino acid residue substituted at position 19a of the model **a** peptide sequence (Wagschall *et al.* 1999) or position 22 in the model **d** peptide sequence (this study).

b $\Delta\Delta G_u$(Ala) is the difference in the free energy of unfolding (ΔG_u) relative to the Ala substituted peptide analog.

c Normalized stability (NS) represents the stability of each substituted analog relative to Gly = 0 and Leu = 100. Values were calculated from the equation: NS = ($\Delta\Delta G_u$(Ala)X + $\Delta\Delta G_u$(Ala)Gly/ $\Delta\Delta G_u$(Ala)Leu + $\Delta\Delta G_u$(Ala)Gly) x 100, were X denotes the stability of any peptide anaolog.

d Normalized occurrence was calculated as (f_X/f_{Leu}) x 100, where f_X is the relative frequency of each amino acid occurring in either an **a** or **d** position of the coiled-coil repeat, and f_{Leu} is the relative frequency of Leu occuring in that same position. Relative frequency represents the frequency of occurrence of each amino acid in that position divided by the overall frequency of occurrence of each amino acid in the GenBank (Lupas et al. 1991).

for *de novo* design applications, the creation of algorithms designed to predict the occurrence and stability of coiled-coils in naturally occurring peptides and proteins, and our ability to understand better the function/structural role of different residue in coiled-coils.

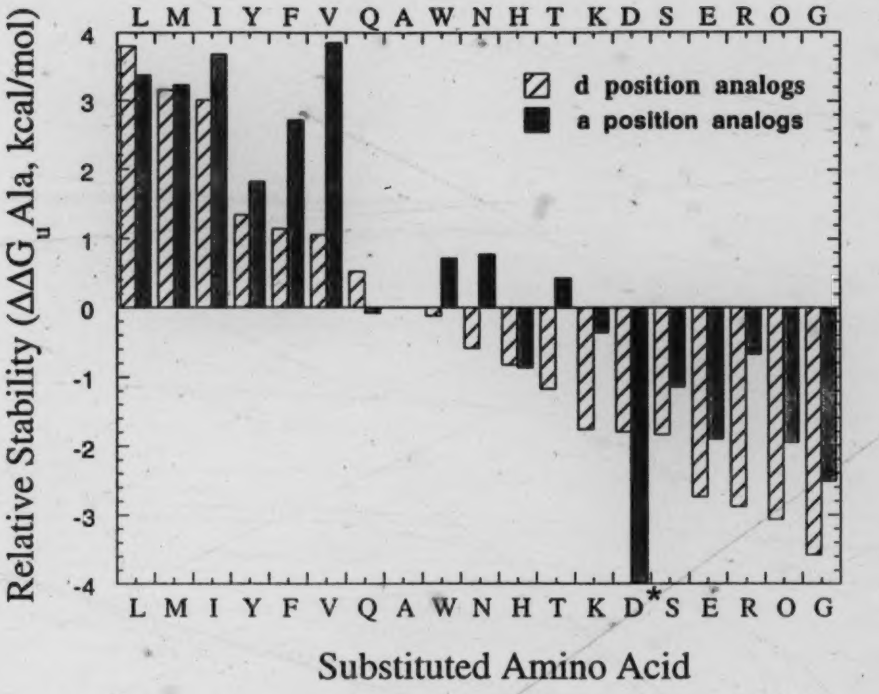
Comparison of position "d" versus "a" stability data.

Comparison of the crystal structures of coiled-coils has revealed two distinct types of packing arrangements within the **a** and **d** positions of two stranded coiled-coils. In the case of the **a** position, the C_α-C_β vector projects out of the interior of the core roughly parallel to the C_α-C_α vector and thus is termed "parallel packing". In contrast, at a **d** position, the C_α-C_β vector projects roughly perpendicular to the C_α-C_α vector and thus is termed "perpendicular packing". As a result, the differences in the frequency of occurrence and stability of certain residues has been ascribed to their ability to pack better in each of these positions. For example, Harbury *et al.*, (1993) described that parallel packing arrangements in the **a** position are more favorable for the β-branched residues (e.g., Ile, Val and Thr), where the methyl group projecting off of the β-carbon is able to pack back into the centre of the core. In contrast, perpendicular packing arrangement observed in a d position are more ideal for leucine residues, where although the γ-carbon moves out of the core, one of the δ-carbons can move back into the core, resulting in excellent geometric complementarity and packing between the symmetry-related side-chains.

To compare and contrast the differences between the $\Delta\Delta G_{u(Ala)}$ stability values obtained in this study with those obtained from the model **a** study (Wagschal *et al.*, 1999b), and thus highlight more clearly the thermodynamic preferences of the individual residues in either an **a** or **d** position, a histogram plot depicting the relative stability values for each data set was plotted. As shown in Figure VIII-9A, large differences in stability for the individual residues in the two positions are observed, and thus allude to positional stability preferences. For example, the β-branched residues of Ile, Val and Thr are clearly more favorable in an **a** position. Leu is more stable in the **d** position. In the case of the charged

Figure VIII-9. Panel A); Comparison of the relative stability contribution each amino acid imparts to the hydrophobic core of a two stranded coiled-coil in either an **a** or **d** position. Relative stability values are as those calculated in Table VIII-1. The Asp **a** position analog (denoted with an asterisk) stability is estimated to be greater than - 4 kcal/mol. Panel B; difference in stability between the **a** and **d** positions. Positive values represent greater stability in an **a** position relative to that in its **d** position. Correspondingly, negative values represent greater stability in a **d** position relative to that observed in an **a** position. The Asp **d**-position analog (denoted with an asterisk) is estimated to be 3 kcal/mol or greater in stability compared to the **a** position analog.

A


L M I Y F V Q A W N H T K D S E R O G
Relative Stability (ΔΔG$_u$ Ala, kcal/mol)
4
3
2
1
0
-1
-2
-3
-4
▨ d position analogs
■ a position analogs
L M I Y F V Q A W N H T K D* S E R O G
Substituted Amino Acid

B

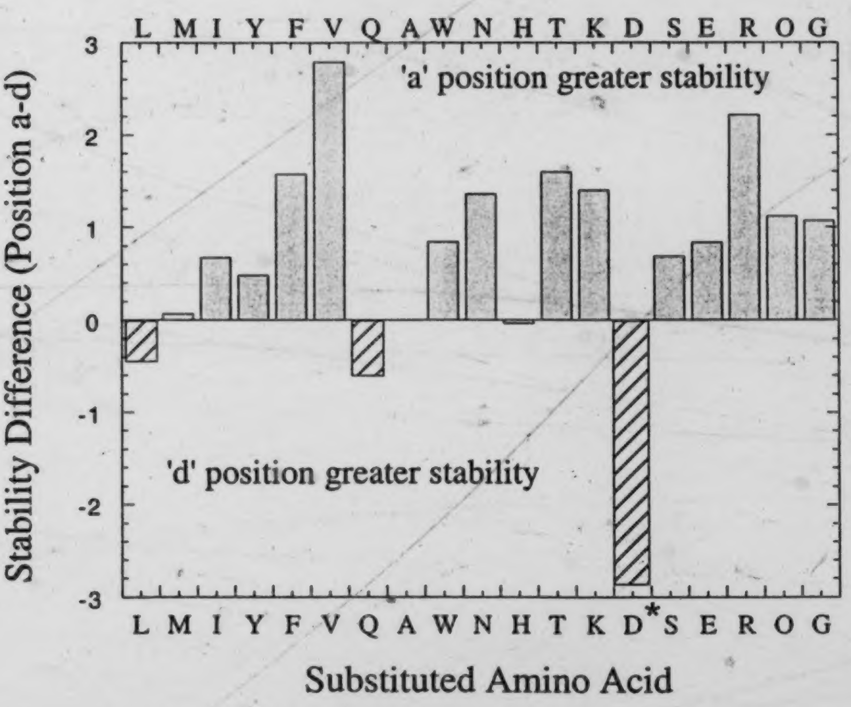

L M I Y F V Q A W N H T K D S E R O G
Stability Difference (Position a-d)
3
2
1
0
-1
-2
-3
'a' position greater stability
'd' position greater stability
L M I Y F V Q A W N H T K D* S E R O G
Substituted Amino Acid

Table VIII-4. Estimation of packing effects for amino acid side-chains in the "a" and "d" positions of a two-stranded α-helical coiled-coil

Amino Acid	Stability Observed[a] "a"	"d"	Positional Difference[b] "a"-"d"	Helix Propensity[c]	Hydrophobicity[d]	Packing Effect[e] "a"	"d"
Leu	1.75	1.90	-0.15	-0.15	1.90	0.00	0.15
Ile	1.95	1.50	0.45	-0.37	2.04	0.28	-0.17
Val	2.05	0.55	**1.50**	-0.54	1.24	**1.35**	-0.15
Met	1.70	1.60	0.10	-0.29	1.26	0.73	0.63
Phe	1.50	0.60	0.90	-0.48	2.02	-0.04	-0.94
Tyr	1.09	0.70	0.39	-0.53	0.89	0.74	0.34
Trp	0.40	-0.05	0.45	-0.47	2.65	**-1.78**	**-2.23**
Ala	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Thr	0.10	-0.60	0.70	-0.68	-0.07	0.85	0.15
Ser	-0.65	-0.90	0.25	-0.63	-0.47	0.45	0.20
Qln	-0.05	0.25	-0.30	-0.35	-0.72	**1.02**	**1.32**
Asn	0.45	-0.30	0.75	-0.63	-1.24	**2.32**	**1.57**
His	-0.60	-0.40	-0.20	-0.63	-0.24	0.27	0.47
Arg	-0.40	-1.45	**1.05**	-0.06	-1.79	**1.45**	0.40
Lys	-0.20	-0.90	0.70	-0.26	-1.77	**1.83**	**1.13**
Glu	-1.00	-1.35	0.35	-0.64	-1.29	0.93	0.58
Asp		-0.90		-0.75	-1.47		**1.32**
Gly	-1.25	-1.80	0.55	-0.96	-0.42	0.13	-0.42

Table VIII-4 Legend

---

a $\Delta\Delta G_{Ala}$ (Stability Observed) is the free energy of unfolding for a single amino acid side-chain in the monomer state relative to alanine analog in the **a** and **d** position, $\Delta\Delta G_{Ala}$ = kcal/mol.

b $\Delta\Delta G_u$ (Position difference) is the difference in free energy of unfolding between the **a** and **d** positions.

c $\Delta\Delta G_{Ala}$ (Helix Propensity) is the free energy difference for a single mutation in a solvent exposed position of a monomeric α-helix relative to alanine as determined from CD studies (Zhou *et al.*, 1994).

d $\Delta\Delta G_{Ala}$ (Hydrophobicity) is the change in solvent transfer free energy relative to alanine (Eisenberg and McLachlan, 1986).

e $\Delta\Delta G_{Ala}$ (Packing Effect) is the difference between $\Delta\Delta G_{Ala}$ (Observed) and the sum of $\Delta\Delta G_{Ala}$ (Helix Propensity) and (Hydrophobicity).

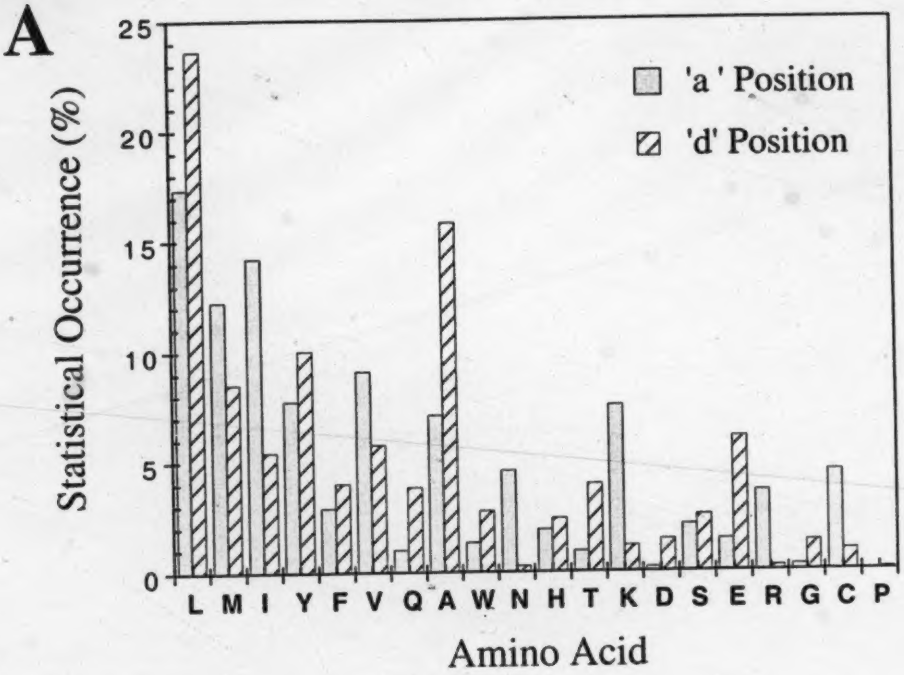
residues, Lys and Arg are clearly more destabilizing in the **d** position compared to that in an **a** position. This is interesting, since both Lys and Arg can initially pack their aliphatic portions to emulate Leu, thus their significantly greater destabilization in the **d** position must reflect the closer distance of the charged group to the hydrophobic core. Other notable differences include Asn and Gln residues. Asn is clearly more stable than Gln in the **a** position, however, this difference is completely reversed in the **d** position. Previously it has been noted how Asn can potentially form an extra hydrogen bond within the core when in an **a** position which likely adds to its stability preference in the **a** position (O'Shea *et al.*, 1991, 1993; Gonzalez *et al.*, 1996a,b,c; Lumb and Kim 1995). Molecular modelling studies have indicated this is not possible in the **d** position and thus the extra methylene group of Gln and greater distance of the amido group from the core is likely responsible for the greater stability of Gln in the **d** position.

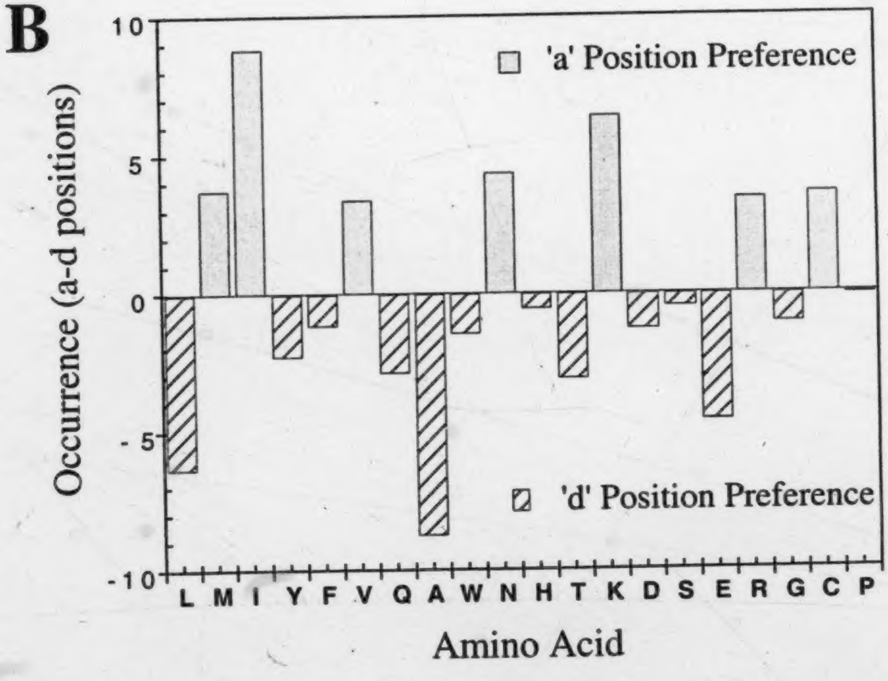
To highlight further the difference between stability of the amino acid residues in an **a** and **d** position, the relative stability of each residue in a **d** position was subtracted from the corresponding relative stability in an **a** position. As shown in Figure VIII-9B and Table VIII-4, almost all residues with the exception of Leu, Gln and Asp were more stable in the **a** position. Thus, these results are in excellent agreement with previous proposals which stated that the packing arrangement at an **a** position, because it directs the side chain initially away from the hydrophobic core, is much more permissive than that in the **d** position, and hence are not as destabilizing (Hu *et al.*, 1990) . Of clear notable stability preference for the **a** position (>1 kcal/mol) were Phe, Val, Asn, Thr, Lys, and Arg. In the case of the **d** position, Leu and Gln are favored along with Asp which is highly destabilizing in the **a** position at 25 °C.

To see if the observed differences between position **a** and **d** stability data sets are also in accord with statistically acquired frequency occurrence differences observed for each amino acid in an **a** or **d** position, similar comparison plots were also made (Fig. VIII-10, panels A and B). In contrast to the stability data, each of the core positions now showed a more even

---

Figure VIII-10. Top; Comparison of the statistical occurrence of each amino acid residue occurring in an **a** or **d** position of the heptad repeat in native coiled-coil sequences (Lupus *et al.*, 1991). Bottom; difference in occurrence between the two positions. Positive values indicate greater occurrence in an **a** position relative to that in its **d** position. Correspondingly, negative values represent greater occurrence in a **d** position relative to that in an **a** position.


A
25
20
15
10
5
0
Statistical Occurrence (%)
'a ' Position
'd' Position
L M I Y F V Q A W N H T K D S E R G C P
Amino Acid


B
10
5
0
- 5
- 10
Occurrence (a-d positions)
'a' Position Preference
'd' Position Preference
L M I Y F V Q A W N H T K D S E R G C P
Amino Acid

distribution between the number of residues which preferred either position. However, similarities between stability and frequency of occurrence do exist. That is, residues which showed high frequency of occurrence for the **a** position were Met, Ile, Val, Asn, Lys, Arg and Cys (4 of the 6 observed by stability), while Leu, Ala and Glu where more frequent in the **d** position (Leu being the most stable observed herein). Thus, taken together, the combined data indeed show positional stability preferences for particular amino acid residues in either the **a** or **d** position which appear to be in good agreement with those observed by statistical frequency occurring in native protein coiled-coil domains (Cohen and Parry, 1990; Hu *et al.*, 1990; Lupas *et al.*, 1991).

C. Discussion

Secondary structure, hydrophobicity and packing

In this chapter, we have determined the effects of substituting 20 different amino acid residues in a central **d** position upon the secondary structure, stability and oligomerization state of a *de novo* designed model two-stranded parallel coiled-coil protein. Our results show that all residues, with the exception of proline, are able to be incorporated into the central **d** position substitution site. Proline's distinct inability to be incorporated into a coiled-coil structure appears to be completely independent of the hydrophobic core position, as it was also unable to form such a structure when substituted in an **a** position, and thus its indifference is primarily ascribed to its strong α-helical breaking nature. The observation that all of the other 18 naturally occurring amino acid residues could be incorporated into a coiled-coil appears to be in good agreement with that of previously acquired statistical occurrence data for such residues in long filamentous coiled-coil proteins (Cohen and Parry, 1990; Lupas *et al.*, 1991; Hu *et al.*, 1990). Interestingly, the frequency of occurrence for several of these residues, e.g., the polar and charged residues in short coiled-coils/leucine zippers (4-5 heptads) is significantly lower or not observed at all. As observed herein, their lower occurrence does not appear to be due to the fact that such residues cannot be incorporated into coiled-coil structures, but rather their low occurrence can be ascribed to their ability to significantly decrease the stability of such structures. For example, as observed above, several residues (Glu, Gly, Asp, Ser and Lys) have the ability to almost completely eliminate the overall stability of the coiled-coil molecule despite the presence of an otherwise optimally packed hydrophobic core.

In regards to hydrophobicity, we observed the general trend that greater stability of the coiled-coil protein correlated with a net increase in hydrophobicity of the substituted "**d**" interfacial residue. Thus, this result is consistent with previous studies byf Hodges *et al.*, (1990) and Zhou *et al.*, (1992a,b) who demonstrated, using similar replacements of an Ala residue with more hydrophobic residues, that only increases in stability of the coiled-coil protein were observed. Further, these results are also in accord with the systematic substitution studies in the hydrophobic cores of globular proteins (Pace *et al.*, 1996 and references therein).

Although it is generally accepted that the hydrophobic effect is the major driving force involved in folding and stabilizing the three-dimensional structures of proteins, both hydrophobicity and packing of the residue sidechains in the hydrophobic core of a protein can affect protein stability (Kellis *et al.*, 1988, 1989; Matsumura *et al.*, 1988; Lim and Sauer 1989; Sandberg & Terwilliger, 1989; Dill, 1990; Eriksson *et al.*, 1992; Mendel *et al.*, 1992). The exact contribution, however, that the packing effects can play in protein folding and stability is still the subject of much debate (Behe *et al.*, 1991). Thus, it has become essential to determine the relative magnitude of such effects in the hydrophobic interiors of proteins within different folds. In the present study, we have addressed this issue by systematically substituting 20 different amino acid residues in the hydrophobic core **d** position of a model coiled-coil. Correspondingly, we have also carried out the same substitutions in the **a** position (Wagschal *et al.* 1999b). Thus, we can not only observe the magnitude of the packing effect in term of each amino acid residue in two geometrically different environments independent of a hydrophobicity change, but also calculate the relative magnitude of the packing effect in relation to hydrophobicity (see below). Table VIII-4, column 3, shows the net difference in stability of each residue between an **a** and **d** position of a two stranded coiled-coil. As shown, the packing factor can be as great as 1.5 kcal/mol (e.g., Val analog) just between the two geometrically different coiled-coil positions. As the difference in stability between all residues was 3.0 and 3.7 kcal/mol per substitution in an **a** and **d** position, respectively, the packing effect term can thus be almost as great as ~ 1/2 of the net stability.

Additionally, the data allows calculation of the magnitude of the packing effect in each position (based on the premise that stability = helical propensity + hydrophobicity + packing), we can see that the packing term can be as great or greater in net magnitude than the hydrophobicity term itself. For example Trp, is very hydrophobic yet due to its poor packing (sterically) results in no apparent net gain in stability (negative packing effect of 1.78-2.23 kcal/mol). Val, on the other hand, which has good hydrophobicity in conjunction with it packing well in an **a** position can more than compensate for the loss in conformational free energy associated with the side chain when restricted in an α-helix (positive packing effect of 1.35 kcal/mol). Similarly, the ability of Asn to pack well and make hydrogen bonds and maintain its conformational entropy can ~ double its net free energy effect compared to that of the hydrophobic effect contribution (positive packing effect of 2.32 kcal/mol). Thus, packing interactions are indeed important if not more in some cases than the hydrophobic effect.

Comparison of the stability results.

A comparison of the stability values observed herein with those from Vinson's group (Moitra *et al.*, 1997), which recently derived a partial list of 7 residues in the 5th **d** position of the bZIP coiled-coil protein, shows excellent agreement for four of the five comparable residues. For example, Met, Ile, Val, and Ser analogs showed $\Delta\Delta G_{Ala}$ values of 2.0, 1.65, 1.1 and -0.25 kcal/mol, respectively, compared to 1.6, 1.5, 0.55, and -0.9 observed in this study. As observed by Vinson, the Leu to Ala substitution caused a free energy change of 4.6 kcal/mol, significantly greater than that of 1.9 kcal/mol observed here. It is not entirely clear why this should be the case; however, it should be noted that a second determination of a Leu to Ala substitution by the same group near the termini of the coiled-coil, was 2.0 kcal/mol, the latter result being within experimental error to our result.

In general, the observation of similar values for the other residues despite their completely different contextual nature, i.e., Vinson's substitution site was located between two valines and positions **e** and **g'** contained Glu and Arg residues compared to that described here, suggests that the observed values herein are largely an intrinsic value for that amino acid at a **d** position. That is not to say that the stability of such residues cannot be influenced by the adjacent residues, but rather the present attempt has been designed to derive such values in the absence of such effects, and with the exception of the Asp analog where molecular modeling studies have clearly shown the potential for possible hydrogen bonding between the side chain carboxyl group and the adjacent **e** and **g'** position Gln residues.

Previously, the Hodges group has also measured the energetic contribution for several residues substituted in a **d** position of a model two-stranded coiled-coil based on sequence of tropomyosin (Hodges *et al.*, 1990; Zhou et al., 1992a,b; Zhu et al., 1993). Relevant to the present discussion was the observation that the energy difference for a Leu to Ala substitution in which the disulfide bond was located in a 2**a** position was 1.6 to 1.8 kcal/mol per substitution in an **a** position and 0.8 to 1.05 kcal/mol per substitution in a **d** position. In the reduced state, both positions demonstrated a similar free energy difference of 1.5 kcal/mol per substitution. This indicates that, in the reduced state, the individual strands of coiled-coils can easily rearrange making each **a** and **d** sites similar, as opposed to when they are locked in an oxidized conformational state. A direct comparison of a Leu to Ala substitution observed herein with those values observed previously shows the result to be in good agreement with that observed for the reduced coiled-coil model (1.9 kcal/mol versus 1.5 kcal/mol, respectively). Moreover, a Leu to Ala substitution in the model X19**a** coiled-coil (Wagschal *et al.*, 1999a,b) also demonstrated a free energy change of 1.7 kcal/mol per substitution. Thus, despite the model **d** (as well as the X19**a**) sequence being oxidized, the location of the disulfide bridge inter-spaced by two glycine residues from the coiled-coil region appears to allow full flexibility of each helical strand to rearrange during substitution. Hence, the values recorded herein are likely to reflect the rearrangement of the two helices minimizing any cavity formation.

It is interesting to note that the energetic difference for a Leu to Ala substitution in the hydrophobic core of the globular protein T4 lysozyme when the cavity size is extrapolated to zero is estimated to be 1.9 kcal/mol per substitution (Erikson *et al.*, 1992), exactly as that observed in this study. In general though, mutations in the hydrophobic core of globular proteins are observed to be much larger than those observed here. For example, Pace *et al.*, (1992) reported values for Ile, Leu, Val, Met, and Phe of 3.8±0.7, 3.5±0.1.1, 2.5±0.9, 3.0±0.9 and 3.8±0.5, kcal/mol, respectively, which are approximately twice the values reported here (Table VIII-1). This difference is most likely explained by the simplicity and ease of the rearrangement of the polypeptide chains in a coiled-coil to eliminate cavity formation, as compared to the larger and more energetically costly rearrangement of more complex globular folds in proteins. Thus, the present results can be taken as the relative free energy term for each substitution in a globular fold in the absence of the second energy term, specified by Eriksson *et al.*, (1992) which depends on the size of the cavity created by the substitution.

Effect of substitutions on the m-values

There has been much discussion regarding the interpretation of the *m* value obtained from plots of ΔG versus denaturant concentration (for review see Bowler et al., 1993; Green *et al.*, 1992; Shortel & Meeker, 1986, 1989; Shortle *et al.*, 1990; Shortel, 1992, 1993; Myers et al., 1995). In brief, it has been suggested that the *m* value reflects the difference in the amount of denaturant bound to the denatured versus the fully folded protein state, and that the differences in the *m* values within a series of mutants reflects differences in the solvent accessible hydrophobic surface area (ΔASA) that is exposed upon unfolding (Schellman 1978). As observed in this study, and similar other host guest studies, a wide variation in the *m* values was observed for the 20 substitutions (1.54 - 3.34 kcal mol^{-1} M^{-1}). Attempts to correlate these values with the calculated ΔASA (data not shown), however, demonstrate a poor correlation between the two values. This was also the case observed for correlation plots between the *m* values and the individual components, i.e., the nonpolar surface area and polar surface area change. Additionally, correlation attempts between the *m* values of this study and those from the model **a** study also showed poor correlation, further suggesting that these variations are not an intrinsic property of the individual amino acids but are dependent upon its local context. Owing that the 19 mutant analogs appeared to start from the same folded state (based upon there similar ellipticity values) and demonstrated single cooperative denaturation transitions indicative of the absence of intermediates, we would have to conclude that the variation in *m* values is reflective of the level of the unfoldedness in the unfolded state. This conclusion would be in good agreement with the recent studies of Smith *et al.*, (1996) who showed that the only physical feature difference between the high and low slope mutants of Protein G was in their unfolded states. It is rather interesting that despite the similarities in the sequences between the **a** and **d** model peptides, the small contextual difference (i.e., the substitution site between two Leu in the **a** model and between two Ile in the **d** model) is sufficient enough to result in differences in the *m* values.

Oligomerization

The second aim of this study was to investigate the effect of each residue sidechain in a **d** position in its ability to discriminate between a particular oligomerization state (i.e. the ability to form a two-stranded vs. a three-stranded coiled-coil structure). As has been noted previously, structural uniqueness arises as a result of a distinct stability difference between two alternative states, and in the case of coiled-coils this has been directly attributed to more favorable packing effects observed in one or the other associative states.

As observed in this study, Leu which packs well in the **d** position and generates the greatest stability, also primarily favors the monomeric two-stranded state. Thus, the high incidence of this residue in leucine zippers can be rationalized by its ability not only to stabilize such structures, but also maintain the two-stranded nature. This is in direct contrast to when it is placed in an **a** position which has been shown to induce only trimer formation.

As expected, based on the packing constraints placed on β-branched residues when existing in the **d** position, Ile, Val and Thr residues clearly promoted trimer formation. Thus, the "acute" packing arrangement in the trimer state is significantly more favorable for these residues than the destabilized two-stranded state ("acute" packing refers to geometries intermediate to those described as parallel and perpendicular). Thus these results appear to be in good agreement with native trimer coiled-coil sequences which show a higher frequency of occurrence of Ile and Val residues in the **d** position, and the mutant GCN4 construct which contained an all Ile hydrophobic core (**a** and **d**) and formed only a trimeric state (Harbury *et al.*, 1993).

Interestingly, Asn which has been noted to play a significant role in defining the two-stranded oligomerization state of leucine zippers in the **a** position (Gonzales *et al.*, 1996a,b,c), appears to demonstrate almost the opposite effect in the **d** position (e.g., 72%:28% trimer:monomer). Thus, as with Thr (above), polarity *per se* is not the sole determinant in defining and/or promoting a specific oligomeric state.

Comparison of the **a** and **d** positions data also reveals how subtle differences such as a single hydroxyl group on the benzyl side-chain group of Tyr can impart a significant effect on the association state. For example, Tyr in the **a** position demonstrated a clear distinction for the trimeric three-stranded state while, in the **d** position, defined only the monomeric two-stranded state. Molecular modeling studies (Lavigne *et al.*, 1999; in preparation) have shown that this unique discrimination ability from that of Phe arises from the ability of Tyr to hydrogen bond via the hydroxyl group to the backbone in the **a** position trimer, increasing its stability while the hydroxyl group sterically clashes in the trimer in the **d** position, reducing its stability from that of the two-stranded state. Interestingly, Tyr occurs twice as often in both the **a** or **d** position than Phe, perhaps due to its ability to help define the unique oligomerization state.

Of the remaining residues, four residues (Gly, Ala, Phe, Ser) show the distinct feature of being largely indiscriminate in defining a specific oligomerization state in both the **a** and **d** positions. These observations suggest that these residues do not stabilize nor destabilize either oligomeric state to any significant degree and thus can be accommodated equally well in either state. This may be an important feature for these residues, as it allows nature to modulate overall protein stability using these residues without imparting any significant effect on the oligomerization choice. The charged residues on the other hand can be classified as promoting largely only the monomeric state. The clear distinction for most of these residues, independent of the **a** or **d** position geometry, likely reflects the greater desolvation cost associated with burying a charge group closer to the hydrophobic core in the three-stranded state vs. the two-stranded state.

Taken together, these results further support earlier observations which showed that the type and placement of residues within the hydrophobic core **a** and **d** positions can indeed significantly influence the final associative state of a coiled-coil structure.

D. Conclusion

In summary, we have investigated the effects of substituting 20 different amino acids residues in a "**d**" position of a *de novo* designed two/three-stranded model coiled-coil. Our results indicate that, although hydrophobicity is a dominant factor in stabilizing these structures, complementary side chain packing is also just as significant a factor. The ability of the model sequence to adopt both a two-stranded and three-stranded structure has also allowed us to observe which residues within a **d** position can impart an effect on the final oligomerization state. The above results clearly show that the β-branched residues of Ile, Val and Thr preferentially promote trimer formation, while the charged residues promote two-stranded monomers. Finally, summation of both of the data sets, substitution in the "**a**" and "**d**" positions, now allows us to form a table of "relative" stability values for each position which can be used to aid in the future *de novo* design of new coiled-coil structures, further investigate the structure function relationships of native coiled-coils, as well as investigate and understand better protein folding issues.

E. References

Bowler, B.E., May, K., Zaragoza, T., York, P., Dong, A., and Caughey, W.S., (1993). *Biochemistry* **32**, 183-190.

Chen, Y.-H., Yang, J. T. and Chau, K. H. (1974). *Biochemistry* **13**, 3350-3359.

Cooper, T. M. and Woody, R. W. (1990). *Biopolymers* **30**, 657-676.

Cohen, C. and Parry, D. A. D. (1990). *Proteins: Struct. Funct. Genet.* **7**, 1-15.

Crick, F. H. C. (1953). *Acta Crystallogr.* **6**, 689-698.

Dill, K.A. (1990) *Biochemistry* **29**,7133-7155.

Eisenberg, D. and McLachlan, A. D. (1986). *Nature* **319**, 199-203.

Engel, M., Williams, R. W. and Erickson, B. W. (1991). *Biochemistry* **30**, 3161-3169.

Eriksson, A. E., Baase, W. A., Zhang, X.-J., Heinz, D. W., Blaber, M., Baldwin, E. P. and Matthews, B. W. (1992). *Science* **255**, 178-183.

Gonzalez, L., Jr., Brown, R. A., Richardson, D. and Alber, T. (1996a). *Nature Struct. Biol.* **3**, 1002-1009.

Gonzalez, L., Jr., Plecs, J. J. and Alber, T. (1996b). *Nature Struct. Biol.* **3**, 510-515.

Gonzalez, L., Jr., Woolfson, D. N. and Alber, T. (1996c). *Nature Struct. Biol.* **3**, 1011-1018.

Green, S.M., Meeker, A., and Shortle D. (1992). *Biochemistry* **31**, 5717-5728.

Harbury, P. B., Zhang, T., Kim, P. S. and Alber, T. (1993). *Science* **262**, 1401-1407.

Hodges, R. S., Saund, A. K., Chong, P. C. S., St.-Pierre, S. A. and Reid, R. E. (1981). *J. Biol. Chem.* **256**, 1214-1224.

Hodges, R. S., Sodek, J., Smillie, L. B. and Jurasek, L. (1972). *Cold Spring Harbor Symp. Quant. Biol.* **37**, 299-310.

Hodges, R. S., Zhou, N. E., Kay, C. M. and Semchuk, P. D. (1990). *Pept. Res.* **3**, 123-137.

Hu, J.C., O'Shea, E. K., Kim, P.S., and Sauer, R.T. (1990) *Science* **250**, 1400-1403.

Kellis, J.T., Jr., Nyberg, K., and Fersht, A.R. (1989). *Biochemistry* **28**, 4914-4922.

Kellis, J.T., Jr., Nyberg, K., Sali, D. and Fersht, A.R. (1988). *Nature* **333**, 784-786.

Lau, S. Y. M., Taneja, A. K. and Hodges, R. S. (1984). *J. Biol. Chem.* **259**, 13253-13261.

Lau, S. Y. M., Taneja, A. K. and Hodges, R. S. (1984). *J. Biol. Chem.* **259**, 13253-13261.

Lim, W.A. and Sauer, R.T. (1989). *Nature* **339**, 31-36.

Lumb, K. J. and Kim, P. S. (1995). *Biochemistry* **34**, 8642-8648.

Lupas, A. (1997). *Curr. Opinion Struct. Biol.* **7**, 388-393.

Lupas, A., Van Dyke, M. and Stock, J. (1991). *Science* **252**, 1162-1164.

McLachlan, A. D. and Stewart, M. (1975). *J Mol Biol* **98**, 293-304.

Matsumura, M., Becktel, W.J., and Mathews, B.W. (1988). *Nature* **334**, 406-410.

Mendel, D., Ellman, J.A., Chand, Z., Veenstra, D.L., Kollman, P.A., and Schuktz, P.G. (1992). *Science* **256**, 1798-1802.

Moitra, J., Szilák, L., Krylov, D. and Vinson, C. (1997). *Biochemistry* **36**, 12567-12573.

Monera, O. D., Sönnichsen, F. D., Hicks, L., Kay, C. M. and Hodges, R. S. (1996). *Protein Eng.* **9**, 353-363.

Monera, O. D., Zhou, N. E., Kay, C. M. and Hodges, R. S. (1993). *J. Biol. Chem.* **268**, 19218-19227.

Monera, O. D., Zhou, N. E., Lavigne, P., Kay, C. M. and Hodges, R. S. (1996). *J Biol Chem* **271**, 3995-4001.

Nelson, M.R. and Chazin, W.J. (1998). **7**, 270-282.

O'Shea, E. K., Klemm, J. D., Kim, P. S. and Alber, T. (1991).*Science* **254**, 539-544.

O'Shea, E. K., Lumb, K. J. and Kim, P. S. (1993). *Curr. Biol.* **3**, 658-667.

Pace, C. N, Shirley, B., Mcnutt, M., and Gajiwala, K. (1996) *FASEB J.* **10**,75-83.

Pace, C.N., Laurents, D.V., and Erickson, R.E (1992) *Biochemistry* **31**, 2728-2734.

Sandberg, W.S.and Terwilliger, T.C.(1989). *Science* **245**, 54-57

Schellman, J.A., (1978). *Biopolymers* **17**, 1305-1322.

Serrano, L. and Fersht, A.R. (1989) *Nature*, 341, 296-299.

Shoemaker, K. R., Kim, P. S., York, E. J., Stewart, J. M. and Baldwin, R. L. (1987). *Nature* **326**, 563-567.

Shortel, D. and Meeker A.K. (1986). *Biochemistry* **28**, 936-944.

Shortel, D. Stites, W.E., and Meeker A.K. (1986). *Biochemistry* **29**, 8033-8041.

Shortel, D., and Meeker A.K. (1986). *Proteins: Struct, Funct, Genet* **1**, 81-89.

Shortle, D. (1992). *Q. Rev. Biophys* **25**, 205-250.

Shortle, D. (1993). *Curr Opin Struct Biol* **3**, 66-74.

Smith, C.K., Bu, Z., Anderson, K.S., Sturtevant, J.M., Engleman, D.M. and Regan, L. (1996). *Protein Sci.* **5**, 2009-2019.

Sönnichsen, F. D., Van Eyk, J. E., Hodges, R. S. and Sykes, B. D. (1992). *Biochemistry* **31**, 8790-8798.

Wagschal, K., Tripet, B, Lavigne, P. Mant, C.T. and Hodges, R.S. (1999b) *J. Biol. Chem.* (submitted).

Wagschal, K., Tripet, B. and Hodges, R. S. (1999a).*J. Mol. Biol.* **285**, 785-803.

Zhou, N. E., Kay, C. M. and Hodges, R. S. (1992a). *Biochemistry* **31**, 5739-5746.

Zhou, N. E., Kay, C. M. and Hodges, R. S. (1992b). *J. Biol. Chem.* **267**, 2664-2670.

Zhu, B. Y., Zhou, N. E., Kay, C. M. and Hodges, R. S. (1993). *Protein Sci.* **2**, 383-394.

Zhu, B.-Y., Zhou, N. E., Kay, C. M. and Hodges, R. S. (1993). *Protein Sci.* **2**, 383-394.

CHAPTER IX

Engineering a *de novo* designed coiled-coil heterodimerization domain for rapid detection, purification and characterization of recombinantly expressed peptides and proteins

A version of this chapter has been published: Brian Tripet, Lei Yu, Daisy L. Bautista, Wah Y. Wong, Randall T. Irvin and Robert S. Hodges, *Protein Eng*. 1996, 10, 101-114.

A. Introduction

The study of protein structure and function is often hampered when only limited amounts of peptides or proteins can be obtained from their native source. To circumvent this problem, many researchers have now turned to using genetic engineering techniques to clone cDNA copies of these genes into bacterial or eukaryotic host systems for easier amplification and recovery. Although amplification results in a greater amount of protein, it does not alleviate the necessary, and sometimes arduous steps required for purification of the desired product to homogeneity. To facilitate in the isolation and purification of recombinantly expressed peptides/proteins, many researchers have now also chosen to fuse onto the desired gene sequence (at its 5′ or 3′ end) a DNA sequence that will encode for an affinity purification "handle" or "tag". By subsequent exploitation of the known binding characteristics of the affinity "tag", a more simplified and rapid purification of the recombinantly expressed fusion products can then be carried out.

Several fusion-based affinity purification systems presently exist and have been used to carry out the purification of a wide variety of different protein products. Some of the different types of fusion tags that have been used in these systems include: protein A (Löwenadler *et al.*, 1986), human serum albumin (Nygren *et al.*, 1988), β-galactosidase

(Ullmann, 1984), glutathionine S-transferase (Smith and Johnston, 1988), the maltose and cellulose binding domains (Bedoulle and Duplay, 1988; Greenwood *et al.*, 1989; Guan *et al.*, 1988), the FLAG epitope tag (Hopp *et al.*, 1988), the strep-tag (Schmidt and Skerra, 1993), the calmodulin binding peptide (Carr, 1991; Stofko-Hahn *et al.*, 1992) and the histidine tag (Hochuli *et al.*, 1987 and 1988; Van Dyke *et al.*, 1992). All of these "tags" differ in several respects with regard to their size, binding interactions and affinities.

Although the above list of available affinity purification "tags" is large, no consensus at present has been reached as to which is the best fusion "tag" for all proteins. In fact, this may be an unattainable expectation, due to the large variety of protein products that are now being expressed which also require a wide variety of conditions for extraction or physiological activity, conditions which are sometimes incompatible with certain fusion tags. There are also a number of inherent limitations that a number of these "native based" purification systems possess that also limit the overall general applicability of these systems. Some of these limitations include availability of the immobilized ligand or receptor, solubility, oligomerization state, steric or conformational effects on the folding and activity of the attached peptide/protein, accessibility to designed cleavage sites, stability and affinity. In addition, another major limitation of existing natural ligand/receptor systems has been in the inflexibility or inability to change or modulate in any way the specificity or affinity of the ligand receptor interaction, this inflexibility often resulting in a fusion tag system that can be used for only a single application.

One approach that could be taken towards achieving a more universal fusion tag system, while overcoming some of the above limitations and disadvantages that exist with "native based" fusion tags, is to try and design a purification system *de novo*, where one could have total flexibility and opportunity to modulate, change and introduce desirable features, or eliminate undesirable ones. The ability to carry out such an approach could also have the benefit of inclusion into the final design of an effective and facile method of detecting the fusion tag, creating a much more versatile system.

Thus, in the present study, the potential of designing *de novo* an affinity purification protein tag system from first principles is explored. We have chosen the two stranded α-helical coiled-coil dimerization motif as a model template for fusion tag design. Choice of this motif was based on three major reasons. Firstly, we required a model template that would allow us the freedom and flexibility to design two distinct and separate interacting species (a ligand and a receptor); a heterodimeric two stranded α-helical coiled-coil would meet this requirement with one strand representing the ligand and the other the receptor. Secondly, a vast amount of information is presently available concerning the structural characteristics and forces that control the formation of the two stranded α-helical coiled-coil. Examples include hydrophobic packing between helices (O'Shea *et al.*, 1991; Zhou *et al.*, 1992a, 1992b; Zhu *et al.*, 1993), electrostatic interactions between helices (Graddis *et al.*, 1993; Kohn *et al.*, 1995a and 1995b; Vinson *et al.*, 1993; Zhou *et al.*, 1994a and 1994b), chain orientation (parallel vs. antiparallel)(Monera *et al.*, 1993, 1994, 1996a), oligomerization state (Bentz *et al.*, 1995; Harbury *et al.*, 1994; Monera *et al.*, 1996b, submitted), polypeptide chain length (Lau *et al.*, 1984a; Su *et al.*, 1994), specificity (homodimerization vs. heterodimerization) (Graddis *et al.*, 1993; O'Shea *et al.*, 1992 and 1993; Schuermann *et al.*, 1991; Zhou *et al.*, 1994b; Zhu *et al.*, 1992) as well as stability (Schuermann *et al.*, 1991; Zhou *et al.*, 1992c; Zhu *et al.*, 1993). For recent reviews describing the two-stranded α-helical coiled-coil system in detail, see Adamson *et al.*, 1993; Alber, 1992; Baxevanis and Vinson, 1993; Hodges, 1992, 1996; Zhou *et al.*, 1992c. Thus, the modeling and design of new and novel sequences for our own application is greatly facilitated. Finally, the coiled-coil dimerization domain has been successfully exploited in a number of other novel applications where specific protein dimerization is required. Some of these include receptor-adhesive modular proteins for interaction with integrins or other types of cell surface receptors (Engel *et al.*, 1991); the generation of mini-antibodies by the fusion of a coiled-coil domain to the Fv subunits (Pack and Pluckthun, 1992; Pack *et al.*, 1995); assembly of soluble T-cell receptor (Chang *et al.*, 1994) and IL-2 receptor (Wu *et al.*, 1994, 1995a and 1995b); and the construction of chimeric proteins including transcription factors by the fusion of a coiled-coil dimerization domain to the functional domain of interest (Francis *et al.*, 1995; Granger-Schnarr *et al.*, 1992; Ottermann and Mekalanos, 1995; Schmidt-Dorr *et al.*, 1991; Sollerbrant *et al.*, 1995; Taylor *et al.*, 1991). Thus, the successful use of this motif for the present application also would seem promising.

In this chapter, we report the *de novo* design, engineering and utilization of a heterodimeric coiled-coil dimerization domain for the purpose of affinity purification of recombinantly expressed peptides. We show that the coiled-coil fusion tag and a fusion peptide containing the tag, can be purified to greater than 95% purity in a single step, as well as being easily detected and characterized using this method.

B. Results

General concept

The general concept behind selective dimerization affinity purification and detection is outlined in Figure IX-1. The target peptide or protein is produced as a fusion to one of the coiled-coil heterodimerization strands. A sample (e.g. crude periplasmic or cell lysate) containing recombinantly expressed fusion peptide or protein is then passed through a column containing the second oppositely charged immobilized strand of the coiled-coil heterodimer. Peptides or proteins containing the fusion tag selectively dimerize to the immobilized strand and are retained. After washing to remove non-specifically bound material, the fusion peptide or protein is eluted and the tag removed enzymatically or chemically if necessary. By similar analogy to purification, detection of the fusion products is carried out at any stage in the process by detecting the fusion tag using a free labeled coiled-coil strand in an immunoblot type format.

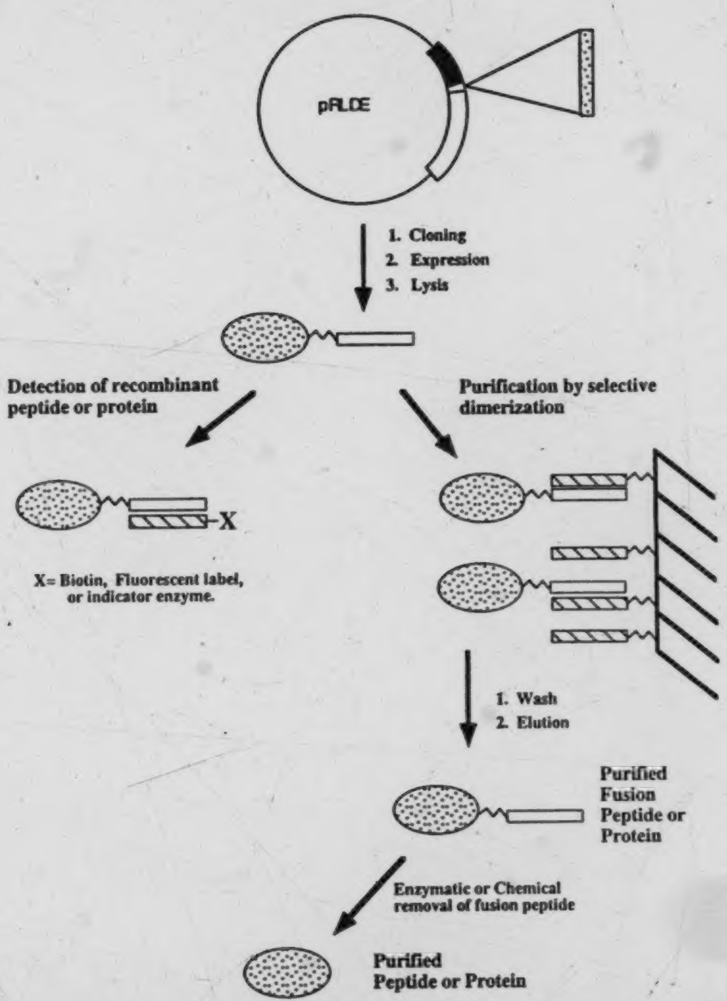

pRLDE
1. Cloning
2. Expression
3. Lysis
Detection of recombinant peptide or protein
Purification by selective dimerization
X
X= Biotin, Fluorescent label, or indicator enzyme.
1. Wash
2. Elution
Purified Fusion Peptide or Protein
Enzymatic or Chemical removal of fusion peptide
Purified Peptide or Protein

Figure IX-1. Schematic representation of the cloning, detection and purification of an E-coil tagged peptide or protein using the coiled-coil fusion system. Top panel shows the vector pRLDE. Protein gene sequences (shaded rectangle) are introduced into the polycloning sites located downstream of the outer membrane protein A signal sequence (solid bar) and upstream of the E-coil gene sequence (open bar). For more details of the expression vector see Figure IX-3. After cloning, recombinantly expressed peptide or protein (shaded oval) is first detected by western blotting using labeled K-coil as a probe (denoted as hatched rectangle attached to X representing biotin or other detection label). The E-coil tag is represented by the open rectangle attached to the recombinant peptide/protein. Purification is carried out by passing cleared periplasmic or whole cell extract from induced cells through a column containing the immobilized K-coil peptide. Peptides or protein containing the E-coil fusion tag selectively dimerize with the immobilized K-coil and are retained. After washing to remove non-specific material, the fusion peptide or protein is eluted. If necessary, chemical or enzymatic cleavage can be carried out to remove the tag.

Table IX-II. Amino acid sequences of synthetic and recombinant peptides.

Peptide Name	Sequence[a]
	|——— Dimerization Domain ———|
	gabcdefg abcdefg abcdefg abcdefg abcdef
K-coil	Ac-KVSALKEKVSALKEKVSALKEKVSALKEKVSALKEGGGnLC*-Amide
E-coil	NH2-ALEGTEFGGGGGGGGEVSALEKEVSALEKEVSALEKEVSALEKEVSALEKGGGGHHHHH-OH
PAK-pilin-E-coil	NH2-ALEGTEFKCTSDQDEQFIPKGCSKFGGGGGGGGEVSALEKEVSALEKEVSALEKEVSALEKEVSALEKGGGGHHHHH-OH
PAK-pilin(met)-E-coil	NH2-ALEGTEFMKCTSDQDEQFIPKGCSKMKFGGGGGGGGEVSALEKEVSALEKEVSALEKEVSALEKEVSALEKGGGGHHHHH-OH

---

a. Residues involved in forming the heterodimeric two stranded α-helical coiled-coil dimerization domain are indicated (|—|). Positions of the heptad repeat are denoted by the letters *abcdefg* , where positions *a* and *d* (bold) represent the 3-4 hydrophobic repeat. The N-terminal amino group and C-terminal carboxyl groups of the synthetically prepared K-coil peptide have been acetylated and amidated. The position of the cysteine used for biotinylation or column immobilization is indicated by the *. The 17 residues of *Pseudomonas aeruginosa* pilin protein , strain K, (PAK-pilin) residues 128-144, are boxed.

De novo design of the coiled-coil sequence

The *de novo* designed coiled-coil sequences used in this study are shown in Table IX-II. The coiled-coil dimerization domain consists of five repeating heptads of the sequence E-V-S-A-L-E-K for the E-coil, and K-V-S-A-L-K-E for the K-coil. Five heptads were chosen as an optimal length for the formation and stability of the coiled-coil domain based on earlier results by Su *et al.* (1994). The hydrophobic residues which create the amphipathic nature of each helix and also drive the folding and stability of the coiled-coil structure were chosen to be valine and leucine in the 'a' and 'd' positions, respectively. The choice of these residues was based on earlier results by Zhu *et al.*(1993) who showed that β-branched residues such as valine and isoleucine are preferred in the 'a' positions, while leucines are preferred in the 'd' positions. In addition, the choice of β-branched amino acids at position 'a' and leucine at 'd' can control the oligomerization state of the coiled-coil (Harbury *et al.*, 1993). The choice of valine over isoleucine was made to optimize the stability of the hydrophobic core so that, in combination with the electrostatic interactions (described below), would ensure preferential heterodimer formation. Alanine was chosen for position 'c' for its high intrinsic helical propensity (Monera *et al.*, 1995; Zhou *et al.*, 1994c). Serine was introduced into position 'b' to increase peptide solubility. Finally, glutamic acid (K-coil) or lysine (E-coil) was introduced into position 'f', the choice determined by the presence of lysine (K-coil) or glutamic acid (E-coil) in positions e' and g'(Figure 2). The introduction of these residues were designed to decrease overall net charge of each strand of the coiled-coil from -10 to -5 (E-coil) or +10 to +5 (K-coil) as well as aid in the overall solubility of the individual strands.

The presence of only glutamic acid or lysine in the 'e' and 'g' positions of the heptad was designed to control the specific and preferential formation of the heterodimer, which is paramount to the usefulness of this domain. It has been shown by a number of earlier studies using native and model coiled-coils (Baxevanis and Vinson, 1993; Graddis *et al.*, 1993; Harbury *et al.* 1993; Monera *et al.*, 1993 and 1994; O'Shea *et al.*, 1991, 1992 and

1993; Schuermann *et al.*, 1991; Vinson *et al.*, 1993; Zhou *et al.*, 1994b) that inter-helical interactions between the 'e' and 'g' positions may not only influence the orientation of the two coiled-coil strands (parallel *vs.* anti-parallel) but also the type of dimerization (homodimerization *vs.* heterodimerization). Several researchers (Graddis *et al.*, 1993; Kohn *et al.*, 1995a and 1995b; Zhou *et al.*, 1994b) have shown that glutamic acid or lysine in the 'e' and 'g' positions adjacent to the hydrophobic interface can create enough charge-charge electrostatic repulsion between two similar strands that coiled-coil homodimer formation is prevented. Figure IX-2 shows helical wheel representations of one heptad of the proposed EE (panel A) and KK (panel C) homodimers. Each heptad creates two pairs of electrostatic repulsions; thus, in a complete dimerization domain of five repeating heptads, ten pairs of interactions will exist and should be sufficient in disrupting any formation of the homodimers, providing the correct choice of residues in the hydrophobic core is also made. In addition, when the two oppositely charged coiled-coil strands (E- and K-coils) attempt to heterodimerize (Figure IX-2, panel B), the charge-charge electrostatic repulsions encountered in the homodimers are replaced by charge-charge electrostatic attractions in the heterodimer, which can now stabilize the heterodimer coiled-coil by ten charge-charge electrostatic attractions. This makes the E-K heterodimer the electrostatically preferred stable structure over the EE or KK coiled-coil homodimers.

In the design of the K-coil strand, the K-coil sequence also contains five extra residues GGG(nL)C (nL=norleucine) attached to its C-terminus. The C-terminal cysteine provides the necessary thiol functional group for ligation to the activated resin beads and biotin for the purification and detection methods. The norleucine was used as an internal standard for amino acid analysis and (static) resin capacity determination, while the three glycine spacer added distance from the column support and conformational flexibility to the immobilized ligand strand.

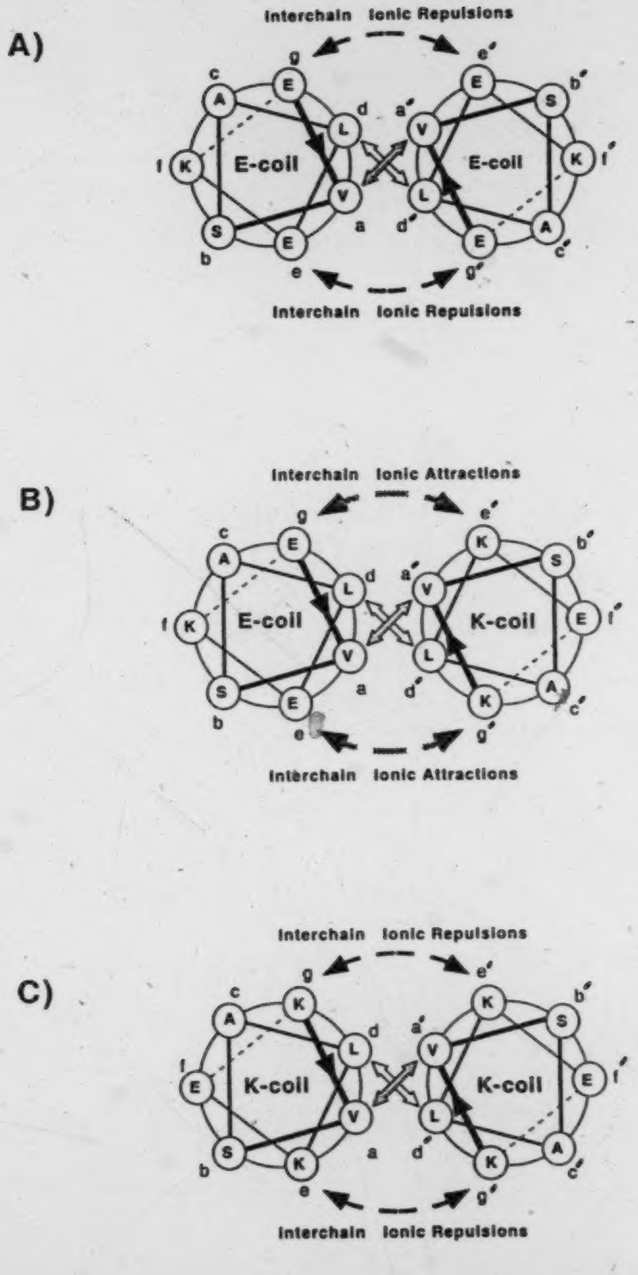

A)
Interchain Ionic Repulsions
g
e'
c
b'
A
E
E
S
d
a'
L
V
f
K
E-coil
E-coil
K
f'
V
L
S
E
E
A
a
d'
b
c'
e
g'
Interchain Ionic Repulsions
B)
Interchain Ionic Attractions
g
e'
c
b'
A
E
K
S
d
a'
L
V
f
K
E-coil
K-coil
E
f'
V
L
S
E
K
A
a
d'
b
c'
e
g'
Interchain Ionic Attractions
C)
Interchain Ionic Repulsions
g
e'
c
b'
A
K
K
S
d
a'
L
V
f
E
K-coil
K-coil
E
f'
V
L
S
K
K
A
a
d'
b
c'
e
g'
Interchain Ionic Repulsions

Figure IX-2. Helical wheel representation of one heptad in three possible coiled-coils (Panel A, EE homodimer; Panel B, EK heterodimer; Panel C, KK homodimer), viewed from the N-terminus. The nomenclature used for the coiled-coils is defined by the amino acid residues in positions e and g, e.g., EE coil contains Glu residues at positions e and g. Wide arrows depict hydrophobic interactions, dashed arrows indicate inter-helical ionic attractions or repulsions depending on the proposed coiled-coil. These interhelical ionic interactions are i to i′+5 (g-e′).

Table IX-I. Sequences of oligonucleotides used in the construction of the expression vector pRLDE, the PAK-pilin, and PAK-pilin(met) genes.

Oligonucleotide$^{\#}$	**Sequence [5'->3']**
1	GGAATTCGGAGGAGGTGGAGGTGGTGGTGGCGAGGTATCCGCTTTAGAG
2	GTCGTATCAAGAGCACTGACCTCTTTTTCGAGAGCAGAAACTTCTTTCTCTAAAGCGGATACCTC
3	GTCGTATCAAGTGCTCTGGAAAAAGAGGTGTCAGCCTTGGAAAAGGAAGTATCAGCACTTGAGAAG
4	CGGGATCCTTATTAGTGATGGTGGTGATGTCCTCCACCGCCCTTCTCAAGTGCTGATAC
5	CTCGAGGGTACCGAATTCCTGCAGAGATCTGCGGCCGCACA
6	AGCTTGTGCGGCCGCAGATCTCTGCAGGAATTCGGTACCCTCGAG
7	AATTCATGAAGTGTACTTCTGACCAAGACGAGCAATTCATCCCTAAGGGTTGTTCCAAAATGA
8	AATTTCATTTTGGAACAACCCTTAGGGATGAATTGCTCGTCTTGGTCAGAAGTACACTTCATG
9	AATTCAAGTGTACTTCTGACCAAGACGAGCAATTCATCCCTAAGGGTTGTTCCA
10	AATTTGGAACAACCCTTAGGGATGAATTGCTCGTCTTGGTCAGAAGTACACTTG

Preferential formation of the heterodimer

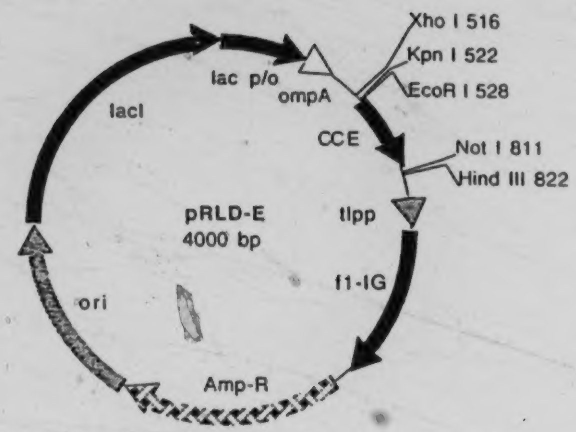
To verify that the two *de novo* designed E and K-coil sequences that constitute the heterodimerization domain will preferentially form a stable heterodimer, as opposed to homodimers, under benign buffer conditions, the secondary structure content of the two synthetic E- and K-coil peptides was investigated by our laboratory (Chao *et al.*, 1996) using CD spectroscopy. In both cases, the individual E- and K-coil peptides show significant random coil structure and only in the presence of each other in a 1:1 ratio do the two peptides show a strong α-helical spectrum with minima at 208 and 222 nm, and a ratio of $[\Theta]_{222}$ *vs.* $[\Theta]_{208}$ of 1.07, indicative of coiled-coil formation. This ratio is consistent with previous observations for coiled-coils (Zhou *et al.*, 1992a and 1992c; Zhu *et al.*, 1992). The CD results show that the electrostatic interactions introduced into the 'e' and 'g' positions are sufficient in disrupting homodimeric formation and directing only the preferential formation of the useful heterodimeric coiled-coil as shown previously for similar peptides (Graddis *et al.*, 1993; Zhou *et al.*, 1994b). Sedimentation equilibrium studies by Chao *et al.* (1996) in our laboratory have shown that these peptides form a very stable heterodimer of the correct molecular weight. Similar sequences to those used in this study were also shown to be heterodimers by size exclusion chromatography and sedimentation equilibrium studies (Graddis *et al.*, 1993; Zhou *et al.*, 1994b). In addition, a kinetic study on the formation of this *de novo* designed heterodimeric coiled-coil, using surface plasmon resonance, was carried out by Chao *et al* (1996). Their biosensor studies showed no evidence for homodimer formation in the concentration range used for the kinetic experiment, and calculated an equilibrium dissociation constant of approximately 1 nM for heterodimer formation.

Construction of the E-coil fusion vector pRLDE

Construction of the *E.coli* expression vector pRLDE was carried out by synthetically preparing and ligating the E-coil gene into the modified vector pRLD (originally pSK40).

Construction of the E-coil gene was carried out by first back translating the *de novo* designed E-coil amino acid sequence (Table IX-II) using a codon bias based on the highly expressed *Pichia pastoris* genes (Koutz *et al.*, 1989). To allow for conformational flexibility and distance of the fusion tag from cloned products, eight glycine codons were also included at the 5' end of the gene. To allow for a second possible method of purification and later comparison of methods, the codons for four glycine and five histidine residues were introduced onto the 3' end of the gene. To avoid the creation of an excessive number of restriction sites within the E-coil gene, the gene was synthesized by preparing only two sets of overlapping oligonucleotides and the remainders filled in by Klenow fragment. The two double-stranded fragments were then digested with Alwn1 I to create sticky ends at one end of each fragment and ligated together with T4 DNA ligase. The final E-coil gene was then digested simultaneously with EcoR1 and BamH1 to generate sticky ends for force-cloning into the plasmid vector.

Modification of the open reading frame and introduction of unique restriction sites into pASK40 was carried out to allow for the easier cloning and transfer of genes between this bacterial expression vector and a eukaryotic *Pichia pastoris* vector also used in this lab. The plasmid pASK40 was first digested with EcoR1 and the overhang ends blunted with Mung-Bean nuclease, shifting the open reading frame and removing the existing EcoR1 site. The blunted plasmid was then digested with Hind III and ligated with a synthetically prepared 43 bp fragment (Table IX-I, oligonucleotides 5 and 6) which contained the new unique restriction sites Xho I, Kpn I and Not I, as well as a new EcoR1 site downstream, creating pRLD. The plasmid pRLD was then digested with EcoR1 and Bgl 1 and force-cloned with the synthetically prepared E-coil gene to produce the modified expression vector pRLDE (Figure IX-3, top). The complete open reading frame and available restriction sites are also shown in Figure IX-3 (bottom). The final modified expression vector pRLDE represents a general bacterial expression vector for the export of



```
ompA-->
ATG AAA AAG ACA GCT ATC GCG ATT GCA GTG GCA CTG GCT GGT TTC
 M   K   K   T   A   I   A   I   A   V   A   L   A   G   F

                          ↓
                               Xho I   Kpn I   Eco RI
GCT ACC GTA GCG CAG GCC GCG CTC GAG GGT ACC GAA TTC GGA GGA
 A   T   V   A   Q   A   A   L   E   G   T   E   F   G   G

                             E-coil-->
GGT GGA GGT GGT GGT GGC GAG GTA TCC GCT TTA GAG AAA GAA GTT
 G   G   G   G   G   G   E   V   S   A   L   E   K   E   V

TCT GCT CTC GAA AAA GAG GTC AGT GCT CTG GAA AAA GAG GTG TCA
 S   A   L   E   K   E   V   S   A   L   E   K   E   V   S

GCC TTG GAA AAG GAA GTA TCA GCA CTT GAG AAG GGC GGT GGA GGA
 A   L   E   K   E   V   S   A   L   E   K   G   G   G   G

                                       Not I       Hind III
CAT CAC CAC CAT CAC TAA TAA GGA TCT GCG GCC GCA CAA GCT TGA
 H   H   H   H   H   *   *
```

Figure IX-3. Top: pRLDE, a 4000 bp general expression vector for bacterial production and purification of peptide and protein fragments. Abbreviations are: *tipp*, lipoprotein transcription terminator; *fl-IG*, filamentous phage replication origin; *Amp-R*, ampicillin resistance gene; *ori*, origin of DNA replication; *lacI*, Lac repressor gene; *lac p/o*, promoter/operator of the lac operon; *omp A*, DNA sequence encoding the leader peptide of *E.coli* outer membrane protein A; *CCE*, DNA sequence encoding the E-coil fusion tag. The unique restriction sites for cloning are indicated. Bottom: the nucleotide and amino acid sequence of the coding strand covering the whole omp A-E-coil gene is shown. Important restriction sites are underlined. The location of the *E.Coli* signal peptidase cleavage site is indicated by the arrow. Termination codons are indicated by *.

recombinantly expressed peptides/proteins into the periplasmic space containing a C-terminal E-coil fusion tag.

Preparation of the affinity matrix

Preparation of the K-coil affinity chromatography matrix used for purification of recombinantly expressed peptides containing the E-coil fusion tag was carried out by immobilizing the oppositely charged K-coil strand of the coiled-coil heterodimer onto a solid silica support. A schematic representation depicting the steps involved in preparing the affinity matrix can be seen in Figure IX-4. Control pore glass (CPG) resin that had been pre-derivatized with the short aminopropyl linker was chosen as the solid support. The choice of silica over other supports was based on the desirable characteristics of silica for high stability under high pressure conditions, small changes in volume with changes in solvents, compatibility with nearly all types of solvents and its previously demonstrated success as a support in high performance affinity chromatography for other immobilized ligands and receptors (Weetall *et al.*, 1974; Scouten, 1981; Mohr *et al.*, 1985; Phillips *et al.*, 1985; Babashak *et al.*, 1988).

The porosity of the glass resin beads was chosen to be 500 Å. This large pore size should effectively allow almost all proteins found within the cell to pass freely through the support while significantly reducing the amount of back pressures required to pump solvent through the matrix.

Activation and reaction of the CPG support with the K-coil ligand was carried out by first activating the resin beads through acetylation of amino groups of the aminopropyl linker with bromoacetic acid. At this point, any remaining free silanol groups which might cause non-specific binding of unrelated cellular components or lead to protein denaturation were capped using trimethylchlorosilane. The activated glass resin beads were then coupled to the synthetically prepared K-coil strand through the side chain sulfhydryl group of the C-terminal cysteine residue. Any remaining bromoacetyl groups which were left un-reacted

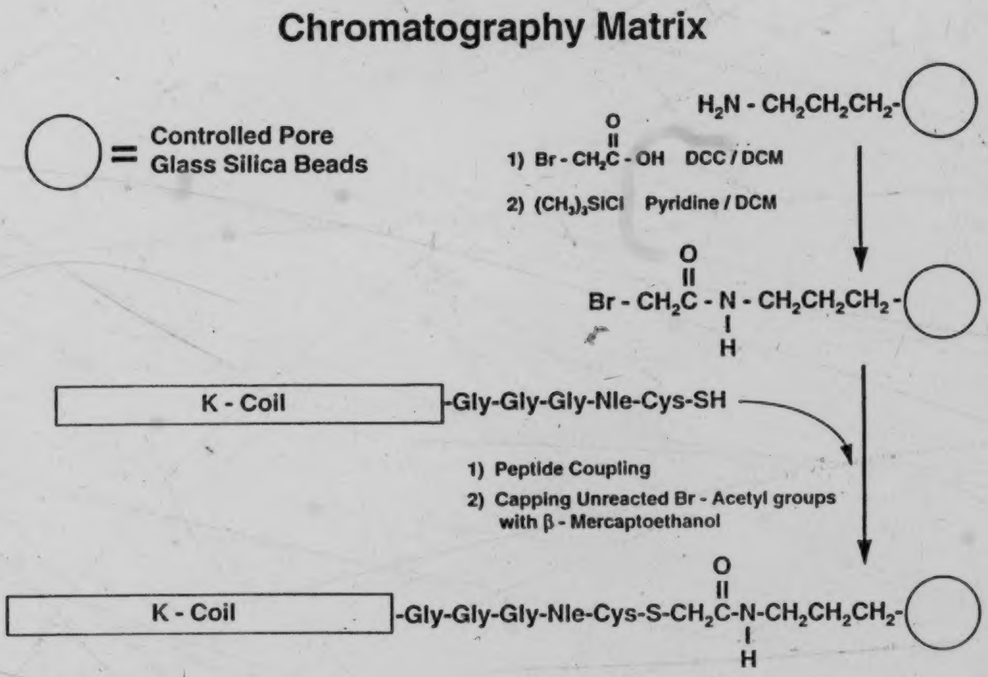

Chromatography Matrix
= Controlled Pore Glass Silica Beads
H₂N - CH₂CH₂CH₂-
1) Br-CH₂C(=O)-OH DCC / DCM
2) (CH₃)₃SiCl Pyridine / DCM
Br - CH₂C(=O) - N(H) - CH₂CH₂CH₂ -
K - Coil
-Gly-Gly-Gly-Nle-Cys-SH
1) Peptide Coupling
2) Capping Unreacted Br - Acetyl groups with β - Mercaptoethanol
K - Coil
-Gly-Gly-Gly-Nle-Cys-S-CH₂C(=O)-N(H)-CH₂CH₂CH₂-

Figure IX-4. Schematic representation of the steps involved in the formation of the K-coil affinity matrix. The control pore glass beads are first acetylated with bromoacetic acid. Remaining free silanol groups are capped with trimethylchlorosilane. Synthetically prepared K-coil peptide containing a C-terminal cysteine residue is coupled to the resin. Quenching of residual bromoacetyl groups is carried out by reaction with excess β-mercaptoethanol, (see Materials and Methods). DCC denotes dicyclohexylcarbodiimide ; DCM denotes dichloromethane.

after coupling were deactivated using β-mercaptoethanol to give a neutral hydrophilic surface.

The combination of the aminopropyl linker with the five terminal spacer residues of the K-coil (GGGnLC) was designed to confer on the K-coil sufficient conformational flexibility and distance from the support to allow for efficient presentation and binding of E-coil fusion tagged peptides/proteins.

Although the initial resin substitution of ~125 μmol/g glass beads could effectively allow for the immobilization of 530 mg of K-coil peptide/g resin, the amount that was coupled was only 12.5 μmol/g resin (50 mg of bound K-coil peptide). The lower resin substitution of 12.5 μmol/g was designed to generate a workable amount of resin that could be easily handled. At 12.5 μmol/g resin, 250 mg of a 20 kD protein could potentially be purified for each gram of resin.

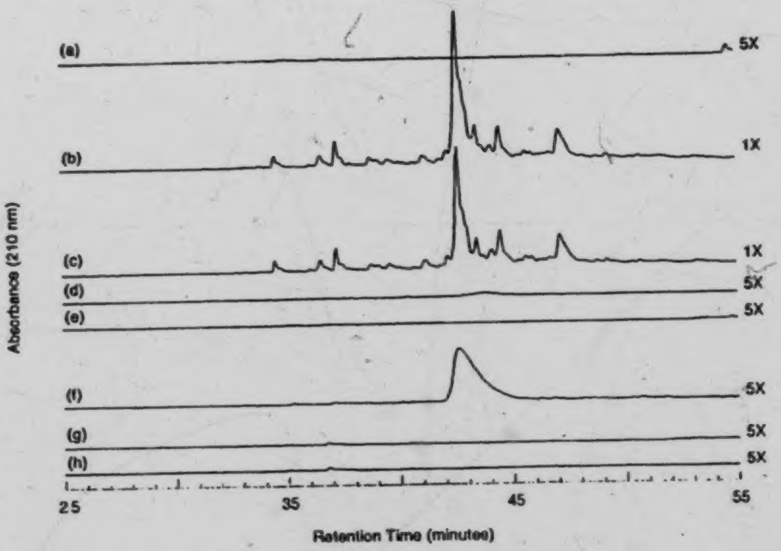
Production and purification of the E-coil and PAK-pilin-E-coil fusion peptide

As examples to demonstrate the general utility and effectiveness of the affinity purification system, we chose to purify and characterize the recombinantly expressed E-coil fusion tag and a 17-residue fusion peptide, PAK-pilin-E-coil (Table II). Transformed JM83 cells expressing the E-coil fusion tag or PAK-pilin-E-coil fusion peptide were osmotically shocked and their periplasmic extracts applied to the immobilized K-coil affinity column for selective dimerization affinity purification. Two wash steps were carried out to remove non-specifically bound materials. An initial 0.5 M KCl wash removed non-specifically bound electrostatic components and a subsequent 80% acetonitrile wash removed organic and lipid components. The RP-HPLC analysis of the eluates eluted from each wash step of the affinity purification method can be seen in Figure IX-5, panels 1 and 2. The crude periplasmic extracts from both peptides show similar chromatograms with the majority of periplasmic extract components being eluted between 35-55 min (Figure IX-5, runs 1b and 2b). After passage through the affinity column, disappearance of the PAK-pilin-E-coil

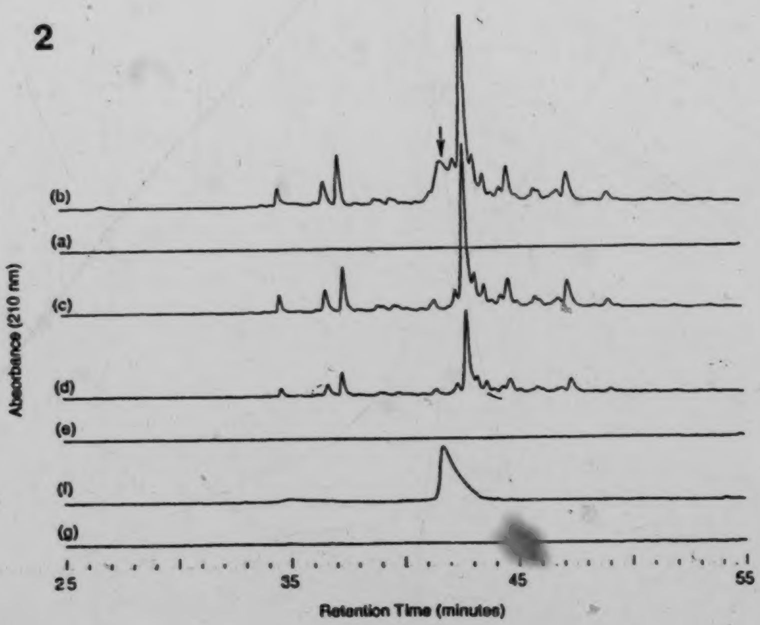
---

Figure IX-5. RP-HPLC analysis of the eluates from each step during affinity purification of recombinantly expressed E-coil and PAK-pilin-E-coil peptides. ***Panel 1***: E-coil purification; (a) pre-elution wash; (b) crude periplasmic extract; (c) breakthrough of periplasmic extract; (d) 0.5 M KCl wash; (e) 80% acetonitrile wash; (f) first elution wash; (g) second elution wash; (h) final elution wash after repassing break-through periplasmic extract through the column. ***Panel 2***: PAK-pilin-E-coil purification; (a) pre-elution wash; (b) crude periplasmic extract; (c) breakthrough of periplasmic extract; (d) 0.5 M KCl wash; (e) 80% acetonitrile wash; (f) first elution wash; (g) second elution wash. All runs were performed on a HP1090 HPLC equipped with a Zorbax RP-C8 column (250 x 4.6 mm I.D., 6.5-μm particle size, 300-Å pore size). Conditions: linear AB gradient (1%B/min), where solvent A is 0.05% TFA in water and solvent B is 0.05% TFA in acetonitrile; flow rate, 1 ml/min; temperature 26°C; detection at 210 nm. Amounts of each sample injected are described in Materials and Methods.

1


(a)
5X
(b)
1X
(c)
1X
(d)
5X
(e)
5X
(f)
5X
(g)
5X
(h)
5X
Absorbance (210 nm)
25
35
45
55
Retention Time (minutes)

2


(b)
(a)
(c)
(d)
(e)
(f)
(g)
Absorbance (210 nm)
25
35
45
55
Retention Time (minutes)

fusion peptide can be easily distinguished from its previous run of complete crude extract as seen in run 2c (location of the fusion peptide indicated by the arrow). In contrast, disappearance of the E-coil fusion tag from crude extract (run 1c) is not detectable due to its smaller peak size and co-elution with much larger and dominant contaminant peaks .

The KCl wash runs (Figure IX-5, runs 1d and 2d) show a greater amount of extract components eluted from the PAK-pilin-E-coil fusion peptide purification than during E-coil tag purification, and thus the amount of non-specific electrostatic binding may be sample dependent. The 80% acetonitrile washes (Figure IX-5, runs 1e and 2e) which do not actually show the elution of any components observable at 210 nm might suggest that this step is unnecessary; however, we have occasionally observed insoluble contaminants in the final products if this step is omitted, and thus routinely include it. Whether these contaminants are residual cell wall components that could have been better removed by a longer centrifugation step or just cellular hydrophobic contaminants that could have been effectively displaced with a lower acetonitrile wash has yet to be investigated.

The observation that the purity of the fusion peptide/protein increases when larger amounts of the E-coil fusion tag is present suggests weak non-specific interactions with the affinity matrix are removed by competitive displacement. This suggests that purification of fusion products could also be optimized by overloading the dynamic capacity of the column.

The combination of both a high ionic strength salt wash and organic wash step to remove non-specifically bound material, while continuing to maintain the stable and specific binding interactions of the affinity purification domain, is a good illustration of one of the unique attributes of this *de novo* designed dimerization domain. Because the coiled-coil heterodimerization domain is composed of, and stabilized by, electrostatic and hydrophobic interactions from the 'a' 'd' 'e' and 'g' positions at the dimer interface (Figure IX-2B), high organic solvent (acetonitrile) washes that would normally break the non-covalent hydrophobic interactions between the two strands (Lau *et al.*, 1984) actually stabilize the coiled-coil dimerization domain by enhancing the electrostatic interactions between the two strands of the coiled-coil in a manner analogous to hydrophilic interaction chromatography (HILIC) (Zhu *et al.*, 1992b). Similarly, high ionic strength salt washes that may disrupt the electrostatic interactions of the domain further enhance the hydrophobic effect that occurs between the two α-helices of the domain in a manner similar to hydrophobic interaction chromatography (HIC)(Ingraham *et al.*, 1985; Mant *et al.*, 1991).

Final elution of the selectively dimerized fusion peptides is carried out by lowering the pH of the elution buffer to pH 2 through the addition of 0.1% TFA and by increasing the acetonitrile concentration to 50%. Low pH causes protonation of the glutamic acid side chains in the 'e' and 'g' positions, thereby disrupting the inter-helical electrostatic e-g´ salt bridges that exist between the two α-helices in benign pH buffer conditions. Although protonation of glutamic acid residues at acidic pH increases their hydrophobicity compared to their ionized state, thus enhancing hydrophobic interactions between the two α-helices by creating a larger hydrophobic interface between the two α-helices (Guo *et al.*, 1986; Hodges *et al.*, 1994; Sereda *et al.*, 1994), the coiled-coil structure can be finally disrupted by increasing the acetonitrile content of the mobile phase to 50% (v/v) to disrupt completely the hydrophobic effect. Chromatograms showing the final affinity purified products can be seen in runs 1f and 2f of Figure IX-5. Integrated areas of the final elution washes show the purity of the final products to be greater then 95%.

One point that should be noted about the eluted peptides is that, although both peptides show unusually broad elution profiles for the final products, this is not a result of neighboring contaminants but rather is an unusual property of the E-coil fusion tag itself in low pH buffer. The same sample can be run at high temperature (80 °C) or pH 6.0 to show a single sharp peak (see Figure IX-6).

To verify that the selectively dimerized peptides retained by the K-coil column are effectively disrupted and eluted during the first elution wash, the column was flushed once again with the final elution solvent and the eluent analyzed for any remaining products. The results of these runs (Figure IX-5, runs 1g and 2g) show no detectable trace of either the E-coil fusion tag or PAK-pilin-E-coil fusion peptide in the second elution wash, indicating that the E/K heterodimer is effectively dissociated and eluted in the first wash.

In order to determine how much of the E-coil fusion tag was missed during initial loading, the breakthrough of the E-coil extract (Figure IX-5, run 1c) was re-loaded onto the affinity column and the complete method repeated. Analysis of the final elution wash (Figure IX-5, run 1h) shows no detectable trace of the E-coil fusion tag, indicating that all of the peptide had been adsorbed onto the affinity column during the first pass.

Purification from whole cell lysate extract

Direction of recombinantly expressed peptide/proteins into the periplasmic space has often been used as a pre-purification step in the final recovery of a homogeneous product. Quite often, however, some expression products are unable to be effectively transported into the periplasmic space or are required to be retained in the cytoplasm. To demonstrate that purification of peptide products containing the E-coil fusion tag can also be carried out just as effectively in the presence of total cellular proteins, *E.coli* cells expressing the E-coil fusion tag and the PAK-pilin-E-coil fusion peptide were grown, induced, and french pressed to liberate all cellular proteins and components. The complete cellular extract was then clarified by filtration, the soluble fraction applied to the K-coil affinity column and the purification method performed. RP-HPLC analysis of the final products eluted after the affinity purification method can be seen in Figure IX-6. The chromatogram of the E-coil fusion tag (Figure IX-6, panel A) does show multiple peaks in the final elution, the largest peak eluted at 23 min being identified as the complete E-coil fusion tag peptide. The smaller peak which is eluted later at 23.5 min was identified by mass spectroscopy to be the E-coil peptide with a double histidine deletion. A possible explanation for the observation of the histidine deletion peptide seen only with the E-coil fusion tag is that multiple clones exist.

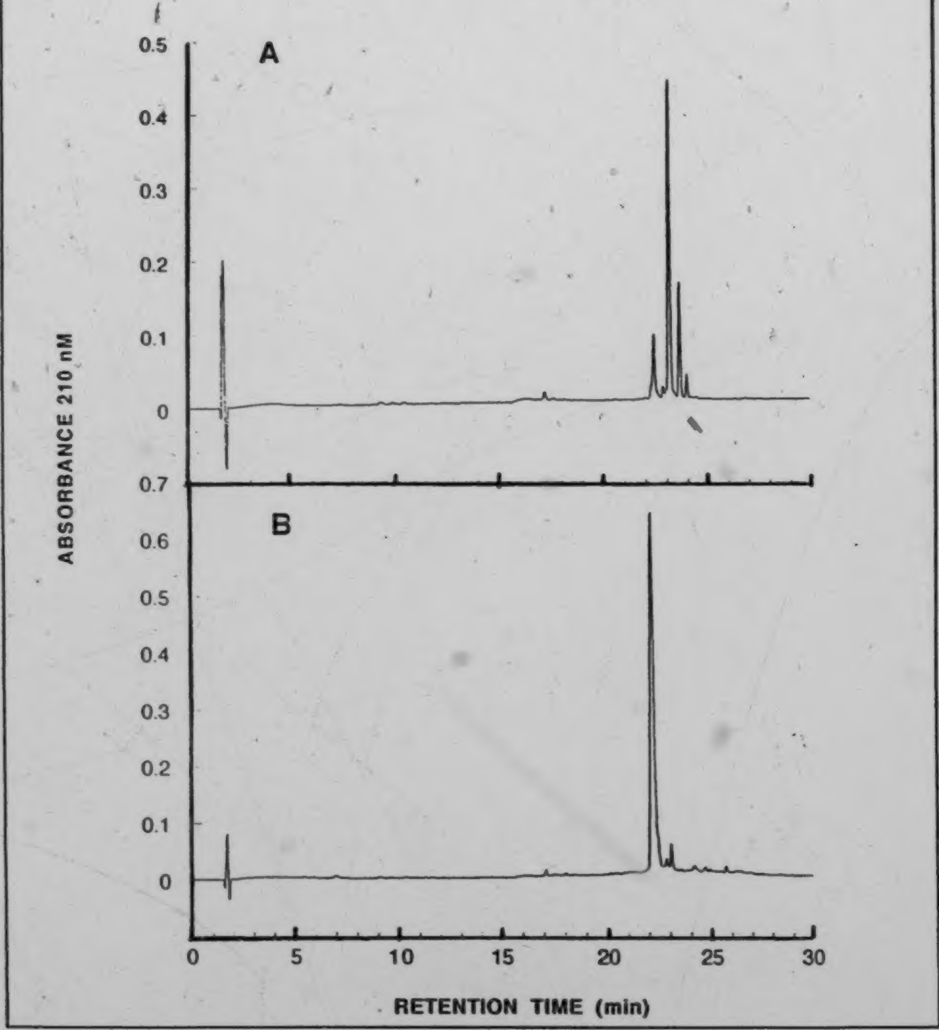


Figure IX-6. RP-HPLC analysis of recovered E-coil and PAK-pilin-E-coil peptides after affinity purification of whole cell extracts. (A) affinity purified E-coil peptide. (B) affinity purified PAK-pilin-E-coil peptide. Conditions: runs were performed on a HP1090 HPLC equipped with a Zorbax SB C-8 column (250 x 4.6 mm I.D., 6.5-µm particle size, 300-Å pore size). A linear AB gradient of 2% B/min for 30 min was employed, where solvent A is 0.05% TFA in water and solvent B is 0.05% TFA in acetonitrile; flow rate, 1 ml/min; temperature, 80°C; detection at 210 nm. Sample loading: 50 µg in 50 µl of 50 mM aq. KH_2PO_4, pH 7.

The chromatogram of the PAK-pilin fusion peptide (Figure IX-6, panel B) shows only a single dominant peak which was identified as PAK-pilin-E-coil peptide by mass spectroscopy, amino acid analysis and western blotting of the product. Figure IX-6, panel B, also shows the excellent single-step purification that can be acheived with this system. It is likely that within the whole cell lysate, native proteins containing coiled-coil sequence exist, however, these proteins do not seem to bind to the K-coil column, either because of the specificity of the preferentially designed interactions for the heterodimer formation or, alternatively, they may bind weakly but are competitively displaced by the E-coil fusion tag. Overall, the results indicate that the purification of E-coil fusion tag and fusion peptide can occur just as effectively from the cellular lysate as from the periplasmic space.

SDS-PAGE analysis of the affinity purified peptides

To analyze the purity of the final affinity purified peptides obtained from either whole cell lysate or crude periplasmic extract, samples of the lysate and extracts before and after affinity purification were analyzed by SDS-PAGE. Electrophoresis was performed on a 15% SDS-PAGE gel and the gel stained with Coomassie Brilliant Blue. Results of the stained gel can be seen in Figure IX-7. The affinity purified PAK-pilin-E-coil peptide from either whole cell lysate or periplasmic extract shows only one band to be present after purification (lanes 4 and 6). The migration of this band correlates with the observed band that appears to be over-expressed in the induced whole cell lysate (lane 3) but not present in uninduced cells (lane 2). Analysis of the crude periplasmic extract (lane 5) also shows that the peptide was effectively exported into the periplasmic space and constitutes a large percentage of the total polypeptides contained in the periplasmic space. Affinity purified E-coil tag peptide recovered from either whole cell lysate or periplasmic extract also shows predominantly one band (lanes 10 and 12). A slightly fainter lower band is also just visible, but this observation correlates with the earlier HPLC chromatogram trace of this product (Figure IX-6) which showed the existence of a second minor peptide peak,

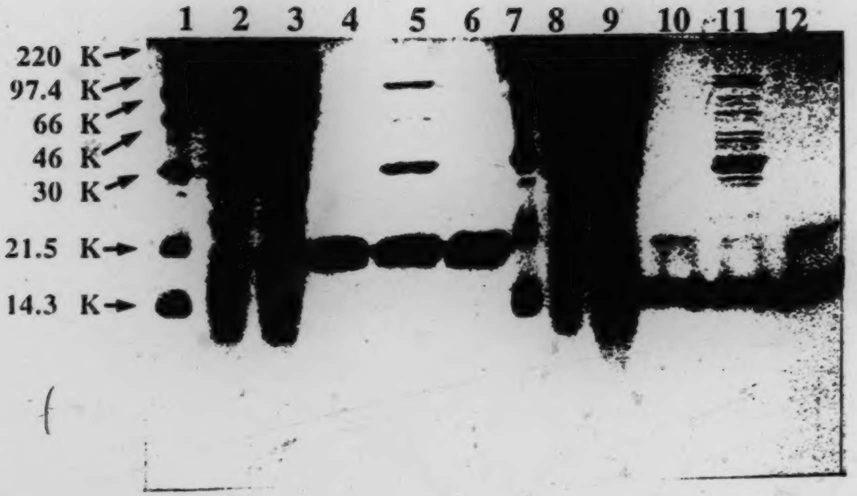


Figure IX-7: SDS-PAGE analysis of the affinity purification of E-coil and PAK-pilin-E-coil peptides from *E.coli* cells. Samples were electrophoresed on a 15% gel and stained with Coomassie Blue. Lane 1, molecular mass markers (kDa); Lane 2, crude PAK-pilin-E-coil whole cell lysate, uninduced (5 μg); Lane 3, crude PAK-pilin-E-coil whole cell lysate,induced (5 μg); Lane 4, affinity purified PAK-pilin-E-coil from induced whole cell lysate (10μg); Lane 5, crude PAK-pilin-E-coil periplasmic extract (10 μg); Lane 6, affinity purified PAK-pilin-E-coil from periplasmic extract (10 μg); Lane 7, molecular mass markers (kDa); Lane 8, crude E-coil whole cell lysate, uninduced (5 μg); Lane 9, crude E-coil whole cell lysate, induced (5 μg); Lane 10, affinity purified E-coil from induced whole cell lysate (20 μg); Lane 11, crude E-coil periplasmic extract (20 μg); Lane 12, affinity purified E-coil from periplasmic extract (20 μg).

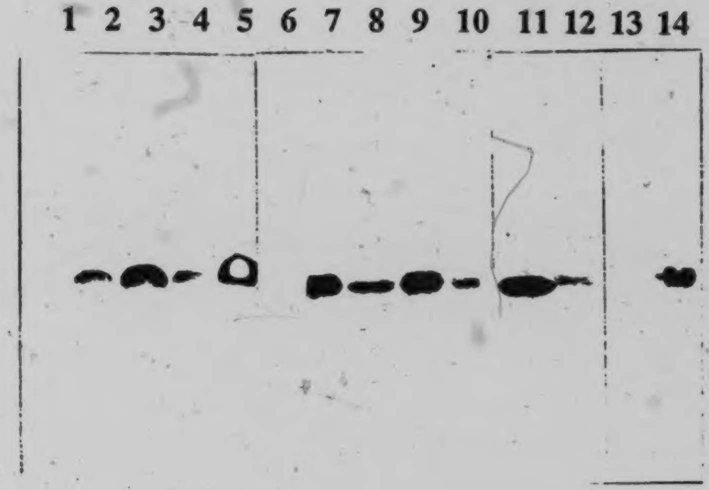


Figure IX-8: Immunodetection of recombinant PAK-pilin-E-coil using PK99H or the biotinylated K-coil. Samples were eletrophoresed on a 15% SDS-PAGE gel and then transferred to a nitrocellulose membranes. Lanes 1-5, 11 and 12 were probed using the mouse monoclonal antibody PK99H (specifically raised against the PAK-pilin protein). Lanes 6-10, 13 and 14 were probed with the biotinylated-K-coil peptide. Lanes 1 and 6, crude PAK-pilin-E-coil whole cell lysate, uninduced (10μg); Lane 2 and 7, crude PAK-pilin-E-coil whole cell lysate, induced (10μg); Lane 3 and 8, affinity purified PAK-pilin-E-coil from induced whole cell lysate (10μg); Lane 4 and 9, crude PAK-pilin-E-coil periplasmic extract (10 μg); Lane 5 and 10, affinity purified PAK-pilin-E-coil from periplasmic extract (10 μg); Lane 11 and 13, PAK-pilin protein (10 μg); Lane 12 and 14, PAK-pilin-E-coil peptide (10 μg).

identified as a double histidine deletion product. The location of both the E-coil tag and the PAK-pilin-E-coil fusion peptide bands on the gel suggests relative molecular weights of ~ 16 kDa and ~20 kDa. These are much larger than the calculated or observed molecular weights by mass spectrometry, and this discrepancy has been attributed to the exceptionally high negative charge density and proportionally fewer SDS molecules bound per unit of mass, thereby retarding its motion and causing it to appear larger than expected.

Detection of the affinity purified PAK-pilin-E-coil peptide using biotinylated K-coil

To demonstrate that the same ligand-receptor interactions used for purification can also be used for detection of the fusion products in an immunoblot type format, samples of PAK-pilin-E-coil fusion peptide whole cell lysate (uninduced, induced and affinity purified), and periplasmic extract (crude, and affinity purified) were loaded onto a 15% SDS-PAGE gel, electrophoresed and transferred to two nitrocellulose membranes. The two membranes were then probed using either the mouse monoclonal antibody PK99H (that specifically recognizes the PAK-pilin-17 mer sequence) or the biotinylated-K-coil peptide. Results of the blots can be seen in Figure IX-8, lanes 1-5. These figures show that the mouse monoclonal antibody PK99H specifically locates and identifies only one band in each sample lane that correlates in migration to the same band that appeared to be over-expressed in the induced whole cell and periplasmic extracts (Figure IX-7, lanes 3 and 5). When the second membrane blot containing the same samples was probed using the biotinylated-K-coil peptide, the K-coil peptide specifically locates and identifies the same protein bands (lanes 6-10) with the same clarity as the antibody. Figure IX-8 (lanes 6 and 7) also shows that the biotinylated-K-coil peptide does not appear to cross-react with any of the other cellular proteins of the prokaryotic host either in the uninduced or induced whole cell lysates, demonstrating probe specificity.

Specificity of the biotinylated-K-coil

To demonstrate that the biotinylated K-coil peptide is specific for the E-coil fusion tag and does not cross-react with the PAK-pilin 17 mer sequence, immunoblots were performed using biotinylated-K-coil peptide and PK99H antibody as primary probes against the PAK-pilin-E-coil fusion peptide and the PAK pilin protein. The results show that the mouse monoclonal antibody PK99H identifies both polypeptides as expected (Figure IX-8, lanes 11 and 12), since both polypeptides contain the epithelial receptor binding domain. However, the biotinylated K-coil peptide probe only specifically identifies the polypeptide sequence which carries the E-coil fusion tag (Figure IX-8, lane 14); in this case, the PAK-pilin-E-coil fusion peptide and not the pilin protein. The overall results indicate that detection with the biotinylated-K-coil peptide is specific and can be as effective as antibodies. Thus, in addition to simplifying and speeding purification of recombinantly expressed peptides or proteins, the coiled-coil fusion tag can also be used for the specific detection of the expressed products prior to purification for screening expressing colonies or after purification for identifying the purified product.

CNBr Digest

Affinity purified recombinant fusion peptides or proteins are frequently desired without the affinity purification "tag" or "handle" still attached. Specific removal of the tag can be carried out by either the inclusion of a specific proteolytic site for proteolytic cleavage or incorporation of specific residues for chemical cleavage. To determine if the E-coil fusion tag could be effectively removed from purified peptide products by chemical cleavage, we designed and expressed the PAK-pilin-E-coil fusion peptide (PAK-pilin(met)-E-coil) to contain two methionine residues next to the desired PAK-pilin 17-mer sequence (Table IX-II). The flanking methionine residues now allow for CNBr specific cleavage sites within the peptide. Using a 50-fold molar excess of CNBr to methionine in 0.1 N HCl, the fusion peptide was quantitatively cleaved from the tag. The fact that the fusion tag contains no methionines and is essentially a random coil allows the reaction to proceed easily with no apparent side reactions or side products generated from the tag itself. Although final purification of the desired PAK-pilin sequence was carried out using RP-HPLC, removal of the tag could have been carried out by neutralizing the cleavage solution and then passing the cleavage mixture down the affinity column once again and collecting the breakthrough.

C. Discussion

Advantages of de novo design

Several advantages of the *de novo* design of an affinity matrix-protein tag system are evident. Firstly, the ability to carry out *de novo* design of an affinity purification system gives one the freedom and flexibility to design a tag system with as many desirable features as possible while attempting to exclude as many undesirable ones. Secondly, the ability to design, *de novo*, a fusion tag system in which the binding interactions and specificity between the ligand and receptor are well understood now allows the opportunity to modulate or change these interactions to suit the desired application. Thirdly, and most importantly, a fusion tag system can be designed to be "multi-purpose", r.e., it can be easily used for purification, detection, immobilization, or in any future application for which the fusion protein products might be employed.

Structural advantages of the coiled-coil heterodimerization domain

The coiled-coil heterodimerization domain designed for this study displays a number of advantageous characteristics that should make it a very useful affinity purification and detection tag system. The small size and random structure (as seen from CD results) of the coiled-coil dimerization tag should eliminate any steric problems associated with the improper folding of proteins attached to the tag as well as ensure the accessibility of enzymes to the tag for enzymatic removal of the tag from peptides or proteins later in the method if desired. The introduction of a large number of charged residues within the fusion peptides sequence generates a tag with good solubility characteristics, a great advantage when attached to weakly or moderately soluble expression products. The designed inability of the E-coil fusion tag to homodimerize (due to the 10 pairs of charge charge electrostatic repulsions in the 'e' and 'g' positions) should also potentially eliminate the problem of insolubility, observed with other fusion tags due to oligomerization and precipitation. The fact that the immobilized ligand strand is small and structureless and can be synthetically or recombinantly prepared allows for large scale or highly substituted columns to be made easily and fairly inexpensively.

Detection methods

A rapid and easy method to detect the affinity purification tag was another important criterion which was considered in the design process. The ability to detect the fusion product specifically, particularly in crude extracts, allows the monitoring of expression levels, assessment of the efficiency of purification protocols, assessment of the stability of the fusion product over time and, of course, identification of the location of the fusion protein from a larger mixture. Many of the previous methods of detection that have been developed to identify fusion proteins by blotting format have relied on the use of specific antibodies. A disadvantage with such an approach include the fact that the application is prevented or hindered when the "fusion tag" antigen used to generate antibodies is unable to do so, or is unable to induce a high titre; frequently, the antibody themselves are unstable, expensive, or cross-reactive. More recently, some fusion systems have now been designed to circumvent some of these problems by using the same receptor/ligand for purification as the probe (by conjugating a detection label onto the free protein itself). The coiled-coil method of detection described in the present study differs from both of the above methods in that it does not require the use of any antibodies or labeled proteins. The K-coil peptide strand of the heterodimer pair is directly linked to the detection of biotin, thereby creating a biotinylated-peptide as the primary probe to detect the presence of the E-coil fusion tag through specific dimerization on the nitrocellulose blots. The advantages of this type of detection is again that it does not require the generation or use of a specific antibody, the biotinylated K-coil peptide can be prepared synthetically (or recombinantly) in large amounts with high purity, the K-coil peptide is a random coil and therefore does not lose its activity (by possibly unfolding) over time and, because of the unique specificity that can be designed into the dimerization domain, cross-reactivity with other proteins should be minimal (as shown with the absence of binding to any cellular proteins with the whole cell extract).

Elution conditions

The choice of the final elution conditions represented one of the more difficult aspects of developing the affinity purification method. Affinity and stability is required to carry out the selective adsorption and retainment, but then the adsorbed material must be released gently and without re-contaminating the purified product with a number of new agents. Several elution strategies exist. One of the most frequently used methods to elute strongly bound substances non-specifically has been by decreasing the pH. More specific benign methods now include competitive elution using free ligands, e.g., iminobiotin (Schmidt and Skerra, 1993), glutathione (Smith and Johnston, 1988), peptide (Brizzard *et al.*, 1994) or maltose (Kellermann *et al.*, 1982). Some newer strategies have even been developed to make the ligand receptor interaction metal-ion dependent; thus, removing the metal ion disrupts the complex and allows elution (Carr, 1991; Hopp *et al.*, 1988; Stofko-Hahn *et al.*, 1992). To be able to incorporate a metal-ion dependent conformational change either through metal ion dependent binding or elution allows for a specific elution that could be carried out in benign pH without the need for a specific natural competing ligand. Such modifications could also be added to future *de novo* designed systems if required.

The final elution conditions used in this study resemble those from RP-HPLC. This solvent mixture, which consists of TFA and 50% acetonitrile, is advantageous for peptide purifications as this solvent can be easily lyophilized, resulting in a freeze-dried (salt and buffer free) product that can be stored or re-suspended in any buffer of choice. Although this solvent is considered "denaturing", its wide-spread use in RP-HPLC for the purification of many proteins (Mant and Hodges, 1991), suggests that the same elution conditions are not a major limitation since most proteins can be easily refolded following removal of the denaturing environment. The use of low pH to elute functionally active proteins has also been well documented, e.g., for the isolation of antibodies. At present, however, a method for benign elution of purified product is being investigated by performing an on-column cleavage with a specific protease and a designed cleavage site. One further point that should be noted about the final elution conditions is that, although the final elution solvents are "denaturing", they pose no particular problems for the immobilized K-coil ligand strand, since this strand reverts to a random coil structure once the second strand of the heterodimer is removed.

Immobilized ligand column

One of the major interests in purifying recombinantly expressed proteins in research is to determine their potential binding interaction with other proteins. This has typically been accomplished by either immobilizing the purified recombinant protein onto a column and then performing affinity chromatography, or pre-incubating fusion proteins with a sample mixture and then co-purifying the bound ligand or receptor with the fusion protein. The opportunity to perform similar studies is also possible with our coiled-coil heterodimer system. Indeed, the ability of this system to co-purify a protein antibody along with the PAK-pilin-E-coil fusion peptide is currently under investigation. The fact that the dimerization domain designed in this study is stable at benign pH, at low pH in the presence of salt, in the presence of acetonitrile and in the presence of high salt concentrations allows the opportunity to explore a variety of different buffer conditions when probing stability and mode of interaction between protein-protein, peptide-protein or ligand receptor interactions.

Other potential applications

Potential applications for the coiled-coil heterodimerization domain other than just for purification and detection also exist. For example, the coiled-coil heterodimerization domain can also be used for the specific immobilization of the recombinantly expressed fusion peptides or proteins onto biosensor chips for the study of their real time binding kinetics. In this case, one of the dimerization strands is specifically immobilized onto the microchip surface and the fusion peptide or protein that contains the second oppositely charged dimerization strand is allowed to interact. The interaction of other molecules to the fusion peptide or protein can then be studied using the biosensor. After analysis, removal of the fusion peptide or protein is carried out by washing the biosensor chip surface with a buffer that disrupts only the heterodimer. Reversion to benign buffer conditions thus allows a second fusion polypeptide to be loaded and analyzed, all in the absence of re-derivatization.

We have already successfully analyzed the binding characteristics of the E and K-coil peptides using this approach (Chao *et al.*, 1996). Results of this study revealed rates of association of $k_{ass} = 1.67\text{-}7.38 \times 10^5\ M^{-1}s^{-1}$, dissociation of $k_{diss} = 2 \times 10^{-4}\ s^{-1}$ and an equilibrium constant $K_d = 0.29\text{-}1.28$ nM. The high rate of binding and slow rate of dissociation agree well with the affinity purification results observed in this study.

Advantages of using the coiled-coil heterodimerization domain in this type of study include the ease with which the ligand can be designed to contain functional groups to allow for easy immobilization onto the biosensor surfaces, as well as easier control over the level of substitution on the surface chip. In addition, fusion products purified using the described method can be used directly for these types of studies. Finally, the coiled-coil domain itself is practically transparent and therefore does not significantly reduce the resonance signal.

Application to the purification of isotopically labeled peptides

Considerable interest has now been directed toward the determination of the 3D structures of ligands when bound to receptors in order to understand better their binding interactions as well as aiding in the design of potential therapeutic agonists or antagonists. To facilitate in the structural process it is often necessary to be able to discriminate clearly between the two sets of signals of the two molecules. In NMR spectroscopy, this "signal overlap" has been largely eliminated by isotopically labeling one of the molecules with N^{15} and C^{13}. Due to the costs involved with such reagents, it would clearly be advantageous to optimize the recovery of labeled product that could be obtained when these products are generated in recombinant systems. In this respect, the coiled-coil heterodimer system should be an ideal system for extracting such labeled polypeptides from the extract. The high apparent binding affinity and selectivity achievable with this method, coupled with the fact that this tag is ideally suited for small molecule products, should ensure successful recovery of products in high yield and purity.

Openness to re-design

Since the coiled-coil heterodimerization domain sequences used in this study result from a basic design in which all of the residue positions and roles are well understood, this allows for the opportunity to re-design the sequences in order to modulate the binding interactions that occur between the two strands of the heterodimer. This contrasts greatly with other available affinity purification systems in which either the exact residues involved in the binding interactions of the ligand receptor complex are not known, or the ligand and receptor are unavailable to be mutated and modified on such an easy scale. This openness to re-design could in fact be used to design coiled-coil sequences with very high stabilities to co-purify and immobilize proteins such as enzymes onto the column to create enzyme bound-columns or likewise be designed to lower their binding affinity strengths to allow for a more benign and "gentler" elution if required. For example, varriations in chain length can be utilized to control affinity in addition to changes in the hydrophobic and electrostatic interactions. Coiled-coils can be formed with as few as three heptads (Su *et al.*, 1994); however, because the determinants of orientation and oligomerization are subtle and varied, one must be careful in changing the sequence.

D. Conclusion

In this study, we have demonstrated that a *de novo* designed coiled coil heterodimerization motif can be used for the rapid detection, purification and characterization of recombinantly expressed peptides. We have specifically shown that the E-coil fusion tag as well as the PAK-pilin fusion peptide containing the E-coil strand of the heterodimer fused onto its C-terminus can be purified to greater than 95% purity in a single step using this method. The purification method described appears to be just as effective in recovering the E-coil fusion tag and peptide from whole cell lysate as that from periplasmic extract and therefore should be appropriate to peptide or protein products expressed into both of these locations. The straight forward detection of E-coil fusion products using a biotinylated K-coil peptide probe has eliminated the need for specific antibodies for routine protein blotting. The CNBr cleavage of the affinity purified PAK-pilin fusion peptide containing the adjacent methionines demonstrates that chemical removal of the fusion tag can be carried out easily and in the absence of any side reactions. The stability of the dimerization domain can also allow immobilized fusion peptide/proteins to be left attached to the affinity column through the non-covalent association of the dimerization domain to generate new ligand bound columns for further studies. Finally, the possibility of re-designing the sequences of the heterodimerization motif to modulate affinity and stability opens the door to successive generations of coiled-coil tags with specific and unique characteristics.

To our knowledge, our results constitute the first report of the use of a *de novo* designed and engineered coiled-coil heterodimerization domain for the purpose of rapid detection, purification, and characterization of recombinantly expressed peptides or proteins.

E. Related Research

Since the reporting of this work, we have sucessfully produced several milligram quantities of N^{15}-labled PAK-Pilin(met)-E-coil peptide. Following CNBr digest, the N^{15}-labled peptides solution structure was determined using heteronuclear multidimensional NMR spectroscopy. This work is reported in Biochemistry 1997, 36, pages 12791-12801 entitled "Solution secondary structure of a bacterially expressed peptide from the receptor binding domain of pseudomonas aeruginosa Pili strain PAK: A heteronuclear multidimensional NMR study" by P.A. Campbell, D.L. Bautista, B. Tripet, W. Wong, R.T. Irvin,R.S. Hodges, and B.D. Sykes.

F. References

Adamson, J. G., Zhou, N. E., and Hodges, R. S. (1993) *Curr. Opin. Biotechnol.*, **4**, 428-437.

Alber, T. (1992) *Curr. Opin. in Genet. & Dev.*, **2**, 205-210.

Ausubel, F., Brent, R., Kingston, R., Moore, D., Seidman, J., Smith, J., and Struhl, K. (1992). Short protocols in Molecular Biology, 2nd Edition: Greene Publishing Associates & John Wiley & Sons.

Babashak, J. K., Phillips, T.M. (1988) *J. Chromatogr.*, **444**, 21-28.

Baxevanis, A. D., and Vinson, C. R. (1993) *Curr. Opin. in Gen. and Dev.*, **3**, 278-285.

Bedoulle, H., and Duplay, P. (1988) *Eur. J. Biochem.*, **171**, 541-549.

Bentz, S., Fairman, R., O'Neil, K., Lear, J., and Degrado, W. (1995) *Phil. Trans. R. Soc. Lond. B*, **348**, 81-88.

Brizzard, B. L., Chubet, R. G., Vizard, D. L. (1994) *BioTechniques* **16**, 730-734.

Carr, D. W. (1991) *J. Biol. Chem.*, **266**, 14188-14192.

Chang, H. C., Bao, Z., Yao, Y., Tse, A. G., Goyarts, E. C., Madsen, M., Kawasaki, E., Brauer, P. P., Sacchettini, J. C., Nathenson, S. G., and Reinherz, E.L. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 11408-11412.

Chao, H., Houston, M., Grothe, S., Kay, C. M., O'Connor-McCourt, M., Irvin, R. T., and Hodges, R. S. (1996) *Biochemistry (in press).*

Craxton, M. (1991) *Methods: A Companion to Methods Enzymol.*, **3**, 20-26.

Engel, M., Williams, R. W., and Erickson, B. W. (1991) *Biochemistry*, **30**, 3161-3169.

Francis, M. K., Phinney, D. G., and Ryder, K. (1995) *J. Biol. Chem.*, **270**, 11502-11513.

Graddis, T. J., Myszka, D. G., and Chaiken, I. M. (1993) *Biochemistry*, **32**, 12664-12671.

Granger-Schnarr, M., Benusiglio, E., Schnarr, M., and Sassone-Corsi, P. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 4236-4239.

Greenwood, J. M., Gilkes, N. R., Kilburn, D. G., Miller, R. C., and Warren, R. A. J. (1989) *FEBS Lett.*, **244**, 127-131.

Gross, E. (1967) *Methods Enzymol.*, **11**, 238-255.

Guan, C. D., Bujard, H., and Bukau, B. (1988) *Gene*, **67**, 21-30.

Guo, D., Mant, C. T., Taneja, A. K., Parker, J. M. R., and Hodges, R. S. (1986) *J. Chromatogr.*, **359**, 499-517.

Harbury, P. B., Zhang,T., Kim, P. S., and Alber, T. (1993) *Science*, **262**, 1401-1407.

Harbury, P. B., Kim, P. S., and Alber, T. (1994) *Nature*, **371**, 80-83.

Hochuli, E., Bannwarth, W., Dölbeli, H., Gentz, R., and Stüber, D. (1988) *Bio/Technology*, **6**, 1321-1325.

Hochuli, E., Dobeli, H., and Schacher, A. (1987) *J. Chromatogr.*, **411**, 177-184.

Hodges, R. S. (1992) *Curr. Biol.*, **2**, 122-124.

Hodges, R. S., Semchuck, P. D., Taneja, A. K., Kay, C. M., Parker, J. M. R., and Mant, C. T. (1988) *Pep. Res.*, **1**, 19-30.

Hodges, R. S., Zhu, B. Y., Zhou, N. E., and Mant, C. T. (1994) *J. Chromatogr.*, **676**, 3-15.

Hodges, R.S., (1996) Biochem. Cell Biol., 74, 133-154.

Hopp, H. P., Prickett, K. S., Price, V. L., Libby, R. T., March, C. J., Cerretti, D. P., Urdal, D. L., and Conlon, P. J. (1988) *Bio/Technology*, **6**, 1204.

Ingraham, R.H., Lau, S.Y.M, Tanjea, A.K., Hodges, R.S., (1985) *J. Chromatogr.*, **327**, 77-92.

Kellermann, O., Ferenci, T. (1982) *Methods Enzymol.* **90**, 459-463.

Kohn, W. D., Kay, C. M., and Hodges, R. S. (1995a) *Prot. Sci.*, **4**, 237-250.

Kohn, W. D., Monera, O. D., Kay, C. M., and Hodges, R. S. (1995b) *J. Biol. Chem.*, **270**, 25495-25506.

Koutz, P., Davis, G. R., Stillman, C., Barringer, K., Cregg, J., and Thill, G. (1989) *Yeast*, **5**, 167-77.

Laemmli, U. K. (1970) *Nature*, **277**, 680-685.

Lau, S. Y. M., Taneja, A. K., and Hodges, R. S. (1984a) *J. Biol. Chem.*, **259**, 13253-13261.

Lau, S. Y. M., Taneja, A. K., and Hodges, R. S. (1984b) *J. Chromatogr.*, **317**, 129-140.

Lee, K. K., Sheth, H. B., Wong, W. Y., Sherburne, R., Paranchych, W., Hodges, R. S., Lingwood, C. A., Krivan, H., and Irvin, R. T. (1994) *Mol. Microbiol.*, **11**, 705-713.

Löwenadler, B., Nilsson, B., Abrahmsén, L., Moks, T., Ljundqvist, L., Holmgren, E., Paleus, S., Josephson, S., Philipson, L., and Uhlén, M. (1986) *EMBO J.*, **5**, 2393-2398.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Journal of Molecular Cloning:A Laboratory Manual, Cold Spring Harbor Laboratory,Cold spring Harbor, New York, NY.

Mant, C.T. and Hodges, R.S. (Eds), (1991) High-Performance Liquid Chromatography of Peptides and Proteins: Seperation, Analysis, and Conformation, CRC Press, Inc.

Maxam, A. M., and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 560-564.

Mohr, P. (1985) Affinity Chromatography: Practical and Theoretical Aspects, Marcel Dekker, New York.

Monera, O. D., Zhou, N. E., Kay, C. M., and Hodges, R. S. (1993) *J. Biol. Chem.*, **268**, 19218-19227.

Monera, O. D., Kay, C. M., and Hodges, R. S. (1994) *Biochemistry*, **33**, 3862-3871.

Monera, O. D., Sereda, T. J., Zhou, N. E., Kay, C. M., and Hodges, R. S. (1995) *J. Pept. Sci.*, **1**, 1-11.

Monera, O. D., Zhou, N. E., Lavigne, P., Kay, C. M., and Hodges, R. S. (1996a) *J. Biol. Chem.*, **271**, 3995-4001.

Monera, O. D., Sonnichsen, F.D., Hicks, L., Kay, C. M., and Hodges, R. S. (1996b) *Protein Engin*, (*in press*).

Nygren, P. A., Eliasson, M., Palmcrantz, E., Abrahmsén, L., and Uhlén, M. (1988) *J. Mol. Recognition*, **1**, 69-74.

O'Shea, E. K., Klemm, J. D., Kim, P. S., and Alber, T. (1991) *Science*, **254**, 539-544.

O'Shea, E. K., Lumb, K. J., and Kim, P. S. (1993) *Curr. Biol.*, **3**, 658-667.

O'Shea, E. K., Rutkowski, R., and Kim, P. S. (1992) *Cell*, **68**, 699-708.

Ottermann, K. M., and Mekalanos, J. J. (1995) *Mol. Microbiol.*, **15**, 719-731.

Pack, P., Muller, K., Zahn, P., and Pluckthun, A. (1995) *J. Mol. Biol.*, **246**, 28-34.

Pack, P., and Pluckthun, A. (1992) *Biochemistry*, **31**, 1579-1584.

Phillips, T. M., Queen, W.D., More, N.S., Thompson, A.M. (1985). *J. Chromatogr.*, **327**: 213-219.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) "Molecular Cloning: A Laboratory Manual", 2nd Edition , Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, NY.

Schmidt, T. G. M., and Skerra, A. (1993) *Protein Eng.*, **6**, 109-122.

Schmidt-Dorr, T., Oertel-Buchheit, P., Pernelle, C., Bracco, L., Schnarr, M., and Granger-Schnarr, M. (1991) *Biochemistry*, **30**, 9657-9664.

Schuermann, M., Hunter, J. B., Hennig, G., and Muller, R. (1991) *Nucleic Acids Res.*, **19**, 739-746.

Scouten, W. H. (1981) Affinity Chromatography: Bioselective Adsorption on Inert Matrices, John Wiley & Sons, New York.

Sereda, T. J., Mant, C. T., Sönnichsen, F. D., and Hodges, R. S. (1994) *J. Chromatogr.*, **676**, 139-153.

Smith, D. B., and Johnston, K. S. (1988) *Gene*, **67**, 31-40.

Sollerbrant, K., Akusjarvi, G., Linder, S., and Svensson, C. (1995) *Nucleic Acids Res.*, **23**, 588-594.

Stofko-Hahn, R. E., Carr, D. W., and Scott, J. D. (1992) *FEBS Lett.*, **302**, 274-278.

Su, J. Y., Hodges, R. S., and Kay, C. M. (1994) *Biochemistry*, **33**, 15501-15510.

Taylor, N., Flemington, E., Kolman, J. K., Bauman, R. P., Speck, S. H., and Miller, G. (1991) *J. Virology*, **65**, 4033-4041.

Towbin, H., Staehelin, T., and Gordon, J. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 4350-4354.

Ullmann, A. (1984) *Gene*, **29**, 27-31.

Van Dyke, M. W., Sirito, M., and Sawadogo, M. (1992) *Gene*, **111**, 99-104.

Vinson, C. R., Hai, T., and Boyd, S. M. (1993) *Genes and Dev.*, **7**, 1047-1058.

Weetall, H. H., Filbert, A.M. (1974). *Methods Enzymol.* **34**, 59-72.

Wu, Z., Eaton, S. F., Laue, T. M., Johnson, K. W., Sana, T. R., and Ciardelli, T. L. (1994) *Protein Eng.*, **7**, 1137-1144.

Wu, Z., Johnson, K. W., Goldstein, B., Choi, Y., Eaton, S. F., Laue, T. M., and Ciardelli, T. L. (1995a) *J. Biol. Chem.*, **270**, 16039-16044.

Wu, Z., Johnson, K. W., and Ciardelli, T. L. (1995b) *J. Immun. Methods*, **183**, 127-130.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) *Gene*, **33**, 103-119.

Zhou, N. E., Kay, C. M., and Hodges, R. S. (1992a) *Biochemistry*, **31**, 5739-5746.

Zhou, N. E., Kay, C. M., and Hodges, R. S. (1992b) *J. Biol. Chem.*, **267**, 2664-2670.

Zhou, N. E., Zhu, B. Y., Kay, C. M., and Hodges, R. S. (1992c) *Biopolymers*, **32**, 419-426.

Zhou, N. E., Kay, C. M., and Hodges, R. S. (1994a) *Protein Eng.*, **7**, 1365-1372.

Zhou, N. E., Kay, C. M., and Hodges, R. S. (1994b) *J. Mol. Biol.*, **237**, 500-512.

Zhou, N. E., Monera, O. D., Kay, C. M., and Hodges, R. S. (1994c) *Prot. Pept. Lett.*, **1**, 114-119.

Zhu, B. Y., Zhou, N. E., Semchuk, P. D., Kay, C. M., and Hodges, R. S. (1992) *Int. J. Pept. Prot. Res.*, **40**, 171-179.

Zhu, B.Y., Mant, C.T., Hodges, R.S. (1992b) *J. Chromatogr.* **594**, 75-86.

Zhu, B. Y., Zhou, N. E., Kay, C. M., and Hodges, R. S. (1993) *Prot. Sci.*, **2**, 383-394.

Chapter X

Future Investigations

TnC-TnI:

Mapping of the distinct sites of interaction between the C-terminal regulatory region of TnI (residues 96-148) and the actin-Tm thin filament and TnC now raises several new and interesting questions . First and foremost, the similarity in sequence observed between the C-terminal region residues 140-148 of TnI with that of the inhibitory region residues 104-115 alludes to the possibility that both of these regions bind similar binding sites on the actin thin filament. This conceivably could be through either two similar binding sites within a single actin monomer unit, or the two binding regions spanning adjacent binding sites within the actin thin filament. To address this question more clearly, we have initiated studies where we have attached a photochemical cross-linking agent (BB) at the N and C-terminal ends of the TnI_{96-148} peptide. Initial experiments (data not shown) show that the BB x-linking group does indeed attach to the actin filament at both locations, however, the exact site(s) of attachment and number of actin units X-linked are still not completely resolved.

Secondly, we have recently combined preparation of the synthetic TnI peptides used in chapter III with the E- and K-coiled-coil heterodimeric display sytem (chapter IX) in order to carrry out surface plasmon resonance kinetic analysis between these regions and TnC. A schematic depicting the various steps are shown in Figure X-1. Results generated to date (data not shown) demonstrate that TnC interacts with both TnI binding regions (TnI_{1-40} and TnI_{96-139}) simultaneously and with high affinity (~$Kd=10^{-8}$ M) in the presence of Ca^{2+}. Moreover, the observed perturbance of a fluoroprobe attached specifically at Cys 98 of TnC by the inhibitory region peptide 96-115 implies that the inhibitory region binding site is not located within the C-domain hydrophobic pocket as initially believed, but rather its binding site is localized to the central helix and the back side of the of the C-domain of

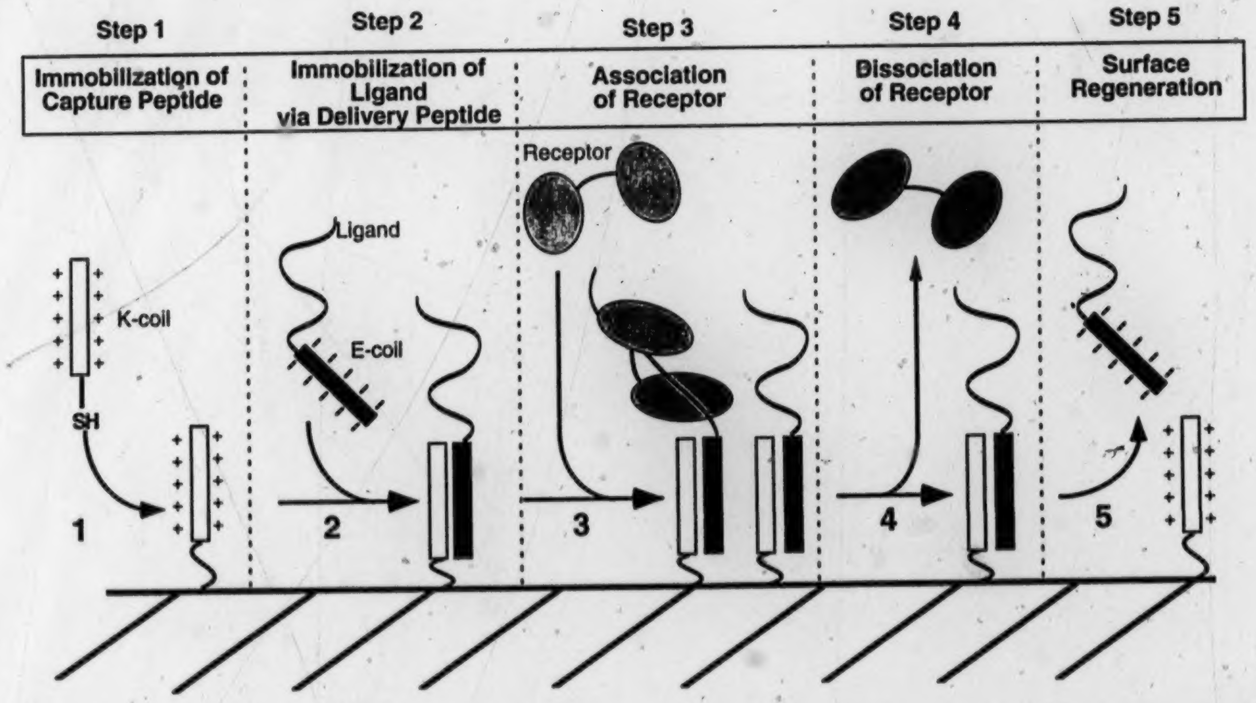

Step 1
Immobilization of Capture Peptide
Step 2
Immobilization of Ligand via Delivery Peptide
Step 3
Association of Receptor
Step 4
Bissociation of Receptor
Step 5
Surface Regeneration
K-coil
SH
1
Ligand
E-coil
2
Receptor
3
4
5

Figure X-1. Schematic representation of the steps involved in utilizing the E- and K-coiled-coil template as a peptide ligand display system for surface plasmon resonance analysis. The sensor chip surface is hatched. The K and E coil peptide strands are denoted as white and black rectangles, respectively. The peptide ligand denoted as a wavy line is conjugated to E-coil. The putative receptor (i.e. TnC) is denoted as a dumbbell shaped molecule (colored in grey). Step 1: the K-coil capture peptide is immobilized to the biosensor surface via its C-terminal cysteine residue using EDC/NHS-PDEA chemistry (see Material and Methods). Step 2: the E-coil-peptide conjugate is loaded (injected) to the desired RU density level . Step 3: the association of receptor to the ligand on cature peptide. Step 4: the dissociation of the receptor from ligand on capture peptide. Step 5: regeneration of surface by ligand removal, allowing the chip surface to be re-used for another ligand-E-coil peptide conjugate.

TnC. These studies are presently on-going and will have to be finished up in order to gain clearer view of the kinetic mechanism of regulatory complex and its change in interactions.

Kinesin:

Structural investigations of the kinesin neck region have demonstrated the existence of a stable two stranded α-helical coiled-coil. Moreover, they showed how two specific helix capping residues can further enhance the stability of this region and possibly play a role in kinesins mechanism of motility. Owing that many other members of the kinesin superfamily are predicted to contain similar coiled-coil regions, future work will be needed to establish the presence of these structures, their thermodynamic stability and if they also have capping residues located at the interface between the coiled-coil region and the beta-linker region.

Coiled-coils:

The results generated from the systematic study of 20 mutations in an "a" and "d" position of a parallel two-stranded model coiled-coil (Chapters VI, VII and VIII) now gives us the necessary data sets to begin several new studies in coiled-coil research.

At the most immediate time will be the further refinement and completion of a predictive algorithm to score, and then identify, coiled-coil regions within native protein sequences based entirely on the physiochemical property to do so. The advantage of this approach over existing methods should be in its ability to predict such regions in the absence of any previously observed statistical frequency bias for specific residues in any particular position and its ability to illuminate unique stability properties within the regions predicted (see chapter V analysis of the kinesin neck region).

The mutational study has also illuminated several residues which can impart either a significant effect upon stability or oligomerization state. We do not know, as of yet, however, the additivity or synergy of each of these effects. Additionally, we do not know the generality of the relative ranked order of stability. That is, to what degree does the context of the site of substitution play in defining the ability of various residues to pack at a particular site.

Even more interestingly, the above data sets now allows us to go back and re-analyze many previously characterized native coiled-coil sequences. These studies should help us to better understand exactly how nature balances the structure and function of these domains, as well as the number of chains in many different protein enviroments. Typically such studies illuminate exception to every rule, which will of course warrant further examinations.

And lastly, the above data sets can now be used to embark on the *de novo* design of many new types of proteins scaffolds, tags and inhibitors of the coiled-coil type.